

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 19, 2003, 09:30:20 : Search time 8148 Seconds

(without alignments)
11417.341 Million cell updates/sec

Title: US-09-494-297-1

Perfect score: 2274
Sequence: 1 atgaaaaaaccaaggttcc.....ggataagaacaatgactag 2274

Scoring table: IDENTITY_NUC
Gap 10.0 , Gape 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenEmbl:*

- 1: gb.ba:*
- 2: gb.htg:*
- 3: gb.in:*
- 4: gb.om:*
- 5: gb.ov:*
- 6: gb.pat:*
- 7: gb.ph:*
- 8: gb.pl:*
- 9: gb.pr:*
- 10: gb.ro:*
- 11: gb.sts:*
- 12: gb.sy:*
- 13: gb.un:*
- 14: gb.vi:*
- 15: em.ba:*
- 16: em.fun:*
- 17: em.hum:*
- 18: em.in:*
- 19: em.mu:*
- 20: em.om:*
- 21: em.or:*
- 22: em.ov:*
- 23: em.pat:*
- 24: em.ph:*
- 25: em.pl:*
- 26: em.ro:*
- 27: em.sts:*
- 28: em.un:*
- 29: em.vi:*
- 30: em.htg.hum:*
- 31: em.htg.in:*
- 32: em.htg.other:*
- 33: em.htg.mu:*
- 34: em.htg.pln:*
- 35: em.htg.rtd:*
- 36: em.htg.mam:*
- 37: em.htg.vrt:*
- 38: em.sy:*
- 39: em.htgo.hum:*
- 40: em.htgo.mu:*
- 41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2274	100.0	11289	1	AE006482	AE006482 Streptococcus
2	2270.8	99.9	2400	1	AB083107	AB083107 Streptococcus
3	756.4	33.3	16207	1	AF447492	AF447492 Streptococcus
4	720.8	31.7	10826	1	SPU49397	SPU49397 Streptococcus
5	664.4	29.2	52900	1	AE014138	AE014138 Streptococcus
6	664.4	29.2	311600	1	AP005141	AP005141 Streptococcus
7	438.4	19.3	11420	1	AE009963	AE009963 Streptococcus
8	215.8	9.5	2845	1	STREPINF	STREPINF Streptococcus
9	180.2	7.9	2263	1	SPSFB	SPSFB Streptococcus
10	167.6	7.4	2066	1	GGU31115	GGU31115 Streptococcus
11	150.2	6.6	2801	1	AF009908	AF009908 Streptococcus
12	97	4.3	321	1	AF009914	AF009914 Streptococcus
13	92.8	4.1	313	1	AF009913	AF009913 Streptococcus
14	75.8	3.3	293650	1	AP004172	AP004172 Streptococcus
15	73	3.2	145265	2	AL935272	AL935272 Streptococcus
16	72	3.2	348	1	AF009920	AF009920 Streptococcus
17	71	3.1	110000	2	PFMAL13_07	PFMAL13_07 Streptococcus
18	69.8	3.1	155204	2	AC007926	AC007926 Streptococcus
19	68.6	3.0	203138	2	BX530070	BX530070 Streptococcus
20	68.4	3.0	309	1	AF009915	AF009915 Streptococcus
21	68.2	3.0	278	1	AF009917	AF009917 Streptococcus
22	68	3.0	250029	3	AE014820	AE014820 Streptococcus
23	67.6	3.0	281	1	AF009909	AF009909 Streptococcus
24	67.6	3.0	282	1	AF009912	AF009912 Streptococcus
25	66.8	2.9	110000	3	AC116984_3	AC116984_3 Streptococcus
26	66.8	2.9	94434	5	AL929250	AL929250 Streptococcus
27	66	2.9	148871	2	AL929469	AL929469 Streptococcus
28	65.6	2.9	213544	2	BX510939	BX510939 Streptococcus
29	65.6	2.9	225581	2	BX537105	BX537105 Streptococcus
30	65.6	2.9	110000	2	BX005010_3	BX005010_3 Streptococcus
31	65	2.9	198550	2	AE004991	AE004991 Streptococcus
32	64.2	2.8	326	1	AF009910	AF009910 Streptococcus
33	64.2	2.8	2642	3	AF103869	AF103869 Streptococcus
34	63.4	2.8	273275	3	AE014828	AE014828 Streptococcus
35	63.4	2.8	330050	3	PFAL929355	PFAL929355 Streptococcus
36	63.4	2.8	250029	3	AE014816	AE014816 Streptococcus
37	62.8	2.7	309	1	AF009918	AF009918 Streptococcus
38	61.6	2.7	185730	2	BX005453	BX005453 Streptococcus
39	61.6	2.7	255	1	AF009916	AF009916 Streptococcus
40	61.4	2.7	110000	2	PFMAL13_24	PFMAL13_24 Streptococcus
41	60.8	2.7	110000	2	PFMAL13_12	PFMAL13_12 Streptococcus
42	60.6	2.7	110000	2	PFMAL13_13	PFMAL13_13 Streptococcus
43	60.6	2.7	333321	3	AC116986	AC116986 Streptococcus
44	60.6	2.7	719	3	AY277504	AY277504 Streptococcus
45	60.4	2.7	719	3	AY277504	AY277504 Streptococcus

ALIGNMENTS

RESULT 1
LOCUS AE006482
DEFINITION Streptococcus pyogenes M1 GAS strain SF370, section II of 167 of the complete genome.
ACCESSION AE006482 AE004092
VERSION AE006482.1 GI:13621422
KEYWORDS
SOURCE
ORGANISM Streptococcus pyogenes M1 GAS
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 11289)
Ferretti J.J., McShan W.M., Adjic D., Savic D., Savic G., Lyon K., Primeaux C., Setate S.S., Surorov A.N., Kenton S., Lai H., Lin S.,

Qian, Y., Jia, H. G., Najjar, F. Z., Ren, Q., Zhu, H., Song, L., White, J.,
Yuan, X., Clifton, S. W., Roe, B. A. and McLaughlin, R. E.
Complete genome sequence of an M1 strain of *Streptococcus pyogenes*
Proc. Natl. Acad. Sci. U.S.A. 98 (8), 4658-4663 (2001)

JOURNAL
MEDLINE
PUBMED
21192684
11296296

AUTHORS
2 (bases 1 to 11289)
Ferretti, J. J., McShan, W. M., Adjee, D., Savic, D., Savic, G., Lyon, K.,
Primeaux, C., Sezate, S. S., Surorov, A. N., Kenton, S., Lai, H., Lin, S.,
Qian, Y., Jia, H. G., Najjar, F. Z., Ren, Q., Zhu, H., Song, L., White, J.,
Yuan, X., Clifton, S. W., Roe, B. A. and McLaughlin, R. E.
Direct Submission
Submitted (10-APR-2001) Department of Microbiology and Immunology,
University of Oklahoma Health Sciences Center, 940 St. Young Blvd,
Oklahoma City, OK 73104, USA

FEATURES
SOURCE
1. 11289
location/qualifiers
/organism="Streptococcus pyogenes M1 GAS"
/mol_type="genomic DNA"
/strain="SF370"
/serotype="M1"
/db_xref="taxon:160490"
complement(1496..2137)
/gene="SPy0121"
complement(1496..2137)
/gene="SPy0121"
/EC_number="2.7.1.113"
/note="Best Blastp hit = dbj|BA016044.1| (AB030884)
similar to *Lactobacillus acidophilus* da/dg-kinase
[*Streptococcus zooepidemicus*]
/codon_start=1
/transl_table=11
/product="putative deoxyguanosine kinase/deoxyadenosine
kinase(I) subunit"
/protein_id="AAK33233.1"
/db_xref="GI:13621423"
/translation="MLIVLAGTIGAGKSSIAALGHELTGDFEYEAVDNPNVLDLYO
DPKRYAFLQIYFLNKRKFSIKAYQADNNILDRSIFEDFLKLNKNVTKTELD
IYQELANLELEGMKRPDLIYIDVSDKMLEIERGRSFEGVDGNPSLEQY
HQVGEPTVYEDYEVSPPMKIDGNSLDFVQNPDLATVLMKIDTKLEHLL"
complement(2157..3134)
/gene="SPy0122"
complement(2157..3134)
/gene="SPy0122"
/note="Best Blastp hit = gb|AAC97155.1| (U49397) unknown
[*Streptococcus pyogenes*]"
/codon_start=1
/transl_table=11
/product="possible transcriptional regulator"
/protein_id="AAK33234.1"
/db_xref="GI:13621424"
/translation="MIKLNSSFRIGVEIPHTVLAPMAGVTSAPFTIATKEGAGLY
VMDISEKGLNNEKTLHMLHIDENHPMSIOFGDAGLKRADFIOINRKADIV
DIMWGPVNVKVEAGAKVLRDPRDKYHLYKVTSTLDPLVTKMRTGADSVLAE
NALAASAGYASALAMHGRPREOMYTGCDHEPTLARVSKATTKIPETGONRVSVOAK
FMLEEIGDAVMTGRAAMNPYLFQINHFETGQELPDLPPAKKIDIAOHKLRLIN
LKGETIAVREFGLAPHYLRGTAGAAKAVRGAVSRAETLAEVAIFETVR"
complement(3121..3993)
/gene="SPy0123"
complement(3121..3993)
/gene="SPy0123"
/note="Best Blastp hit = dbj|BA03806.1| (AP001507)
unknown conserved protein [Bacillus halodurans]"
/codon_start=1
/transl_table=11
/product="putative heat shock protein HSP33"
/protein_id="AAK33235.1"
/db_xref="GI:13621425"
/translation="MDKIISIAQSGAFRAYVLDSTETVALAQEKHNTLSSSTVALGR
TLIANQILANQKGDSKIYKVIQSGDSFGHIIISVADTKGVKGYIONTGVDIKRTATG
EVLVGPFGMGHGVFTIIDGNTSTPLTIGEDFAVYLTSEQCPISGLVY
LIDENKVKVAGGFVOVLPGASEEIAREKRELOMPAISHLASKNHVDLLEIY
GDPEYKRLSEPELSPQDCSREKREFAALMTLPKADQAMIDEDKGAETVQFCGTIKYQ

gene
CDS
FNSDLEAIIISDKA"
complement(4140..5633)
/gene="rofa"
/note="synonym: SPy0124"
complement(4140..5633)
/gene="rofa"
/note="Best Blastp hit = gb|AAC43219.1| (U01312) RoFA
[*Streptococcus pyogenes*]"
/codon_start=1
/transl_table=11
/product="regulatory protein"
/protein_id="AAK33236.1"
/db_xref="GI:13621426"
/translation="MIKELYSIESKCOLIVLPFKTSYLPITEVAEKTGLTFEQLNH
YCEELNAPFGPSLMTIOKRMISQGFHPKRETYLYQVASSVNLQDLAPFLINKSGS
RPLDFAHSPIFNSSAYRMRBMLIPLRMRLEKLSKNTKIVGEERYIRYLIALLSFE
GIKYVDLTQODKNTIHSFLSHSTHLKTSWLSRSEFSYDILALSKRRQFSYTTIQ
TRIFQDLKRLVYDSLSKSSHDIIETQCUNFSAGLDLYLYITANNSFASLOWTP
EHIRQYQLEENDTERLNLNPIITLPLNKEQKASLVKALMFESKSFENLOHFTPE
TNLFVSPYKGNOKLYSLKLIYEEMMAKLPGRDLNKHKHFHLCHVEQSLRNIOFP
LVVFEVANSFTINAHLLTDSPPRYSKSIDFHSYLLQDNVYOIPDLKPOLVITHSOL
IPVYHETKGIQVAELSPDESILSIQELMYQVKEKFOADLRKOL"
8324..8719
/gene="SPy0127"
8324..8719
/note="Best Blastp hit = gb|AAC97149.1| (U49397) LepA
[*Streptococcus pyogenes*]"
/codon_start=1
/transl_table=11
/product="putative signal peptidase I"
/protein_id="AAK33237.1"
/db_xref="GI:13621427"
/translation="MIKRNMDAPSVKAGDAILEPRLSOTYKVEEAVVYEDSSTIK
VGRITIAQGEVDLTGEGELKINGHONBEITFKSREANYPPIRIDNSYLLINDIYS
QESENTLQDAIAKDAIKGTINTLIRLNH"
8752..9774
/gene="SPy0128"
8752..9774
/note="Best Blastp hit = gb|AAC97152.1| (U49397) unknown
[*Streptococcus pyogenes*]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK33238.1"
/db_xref="GI:13621428"
/translation="MKIRHLILTGALTSRAATVGHGVNGAKLTVNKNLDVNSN
ALIPNTDFTKIEPDTTVNEDGKRGKGVALNPTKTYTTINSKGSNRTTAEDFSE
VTEKPEVYIKVTEKIDVPGVSTDTSTYQVAVLMEQQRKFAVYIVGKESG
KVPIDFKNLSIDSTTLVKKRVSGTGDGRSKDFNGLKANQYKASEKVMIEKFG
GQAPVOTKESIDTLLVTKKIDGESIVTNLNPVGVVVEDDYKSEKTYTNVEVSOD
GAVNKIAGNSTEQESTDKDMTITFNKKDFEVPQVAMTVABYIALGIVAVGALYF
VKKKNA"
9784..10497
/gene="SPy0129"
9784..10497
/note="Best Blastp hit = gb|AAC97152.1| (U49397) unknown
[*Streptococcus pyogenes*]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK33239.1"
/db_xref="GI:13621429"
/translation="MTVRLIKLIDKLINIVYICFPFLCIIIALGTYDLATVYOGANA
TNQOYKRRKQVQFDLLAINSVMALTYKGHIDYPIYOGENNNLEINKSVGEISL
SGSVFLDYRNKVFEDKYSLITAHMAGNAREDLNFKRSKSPFNHKEISLETKTKQ
KLRINIFACIOTDAPDSLFPNPIDIVDISKNEELNHIKOKSVQYREILITNTSEFVAL
STCEDMTTGRILVIGQIE"
10519..11166
/gene="SPy0130"
10519..11166

OY 1921 GTTAAACAGATTAACAAACCTCGAATTTAAAGTGTAAAGCAACCATTAATTTAA 1980
|||||
DB 7816 GTTAAACAGATTAACAAACCTCGAATTTAAAGTGTAAAGCAACCATTAATTTAA 7875
OY 1981 CATGGGGAAGTTTAAACACTTCAAGGTTTACCAGAAGTTATCTTACCTGTCAAGAA 2040
|||||
DB 7876 CATGGGGAAGTTTAAACACTTCAAGGTTTACCAGAAGTTATCTTACCTGTCAAGAA 7935
OY 2041 ACAATTTCTGAAGCTATAGGTTAAAGTTAATAGCCAGAAGTAGCAATGCTACAGT 2100
|||||
DB 7936 ACAATTTCTGAAGCTATAGGTTAAAGTTAATAGCCAGAAGTAGCAATGCTACAGT 7995
OY 2101 TCAAAAACAGAAATTAACAGATGATGAGACTGCTTTGAAATTAATTAAGAGCTGT 2160
|||||
DB 7996 TCAAAAACAGAAATTAACAGATGATGAGACTGCTTTGAAATTAATTAAGAGCTGT 8055
OY 2161 GTTCTTACAGAGATTGATCAAAAGATCAATGCTATAGCTTTGATTAATGCTGCTGT 2220
|||||
DB 8056 GTTCTTACAGAGATTGATCAAAAGATCAATGCTATAGCTTTGATTAATGCTGCTGT 8115
OY 2221 ATCAGTTGGGAGTCTGGGGAATTCACACGATTAAGGATAAGAAACATGACTAG 2274
|||||
DB 8116 ATCAGTTGGGAGTCTGGGGAATTCACACGATTAAGGATAAGAAACATGACTAG 8169

RESULT 2
AB083107 2400 bp DNA linear BCT 21-JAN-2003
LOCUS Streptococcus pyogenes cpai gene for collagen binding protein,
DEFINITION complete cds.
ACCESSION AB083107 GI:23503477
VERSION
KEYWORDS
SOURCE
ORGANISM

Streptococcus pyogenes
Streptococcus pyogenes
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

1
Miyoshi-Akiyama,T., Wakisaka,N., Zhao,J. and Uchiyama,T.
Characterization of collagen binding protein (Cpa) of M1-type
group A streptococcus
Unpublished
2 (bases 1 to 2400)
Miyoshi-Akiyama,T., Wakisaka,N., Zhao,J. and Uchiyama,T.
Direct Submission
Submitted (03-APR-2002) Tohru Miyoshi-Akiyama, Tokyo Women's
Medical University, Department of Microbiology and Immunology; 8-1,
Kanda-cho, Shinjuku-ku, Tokyo 162-8666, Japan
(E-mail:KFA01112@nifty.ne.jp, Tel:81-3-3353-8111(ex.22713),
Fax:81-3-5269-7411)
Location/Qualifiers

FEATURES
source

1. 2400
/organism="Streptococcus pyogenes"

/mol_type="genomic DNA"

/strain="291"

/serotype="M1"

/db_xref="taxon:1314"

/country="Japan"

94. 2367

/gene="cpai"

94. 2367

/gene="cpai"

94. 2367

/codon_start=1

/transl_table=11

/product="collagen binding protein"

/protein_id="BAC20340.1"

/db_xref="GI:23503478"

/translation="MKTRPNNKLTNTQVLSKSKRFTVLGVLMFALVTSK
VGAKTVGLVSSSTPNAIIPDSSSEYRGHPYKOPRAVHDLRVNLSGR
SYOYCNKIKKAPLSDSSYKKKKKHDGISTKFEYDAISPRITGELNKLAAVY
NGHPQNAKTEIGLEPLNAIRVTOEAVVYSDNAPISNPDSFRRESLSIVTSLS
LMROALKOLIDPNIATKMPKOVDPDFOLISFESBDDKDKYGNLLSGVLPTKRP

TPDPPMPNPDPOTSVLIRRYAIGDYSKLLEGATLQTDGNVNSFOARFSSNDIGE
RIELSDCTVTLLELNSPAGYSIAPRIEYKAGGVNVLIDGKOLENKNKEIPEYSVE
AYNDPEEFSVLTITONVAKFYAKKNSSQVYKFNVDISPDSEGGCTMPDPTT
GEVATYTHIAGDLEKYVAKPRDIDPDTFLHKIRKVIKGRGQALREYSGLETOLR
AAQDLAIYFEDSDELDDKIKIDHGDNDSTLAVAKILVEYADQSNPDLTDLF
FIPNNKRYOSLIGTQWHEPDLVDIIRHEDKKEVLPVHAKILKRTVYGLAGDRKDFH
FEILKNKNOELISQVTKETNLEFDGKATILKKGESLTLQGLEEGSYLVKED
SEGKRVKNVSEOVNATVSKTGISDSTLAFENNKKEVYVTVGVQDKINGLIALIVAG
ISLQKMGTHIRIRKHD"

BASE COUNT 873 a 364 c 463 g 700 t
ORIGIN

Query Match 99.9%; Score 2270.8; DB 1; Length 2400;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ATGAAAAAACAGGTTTCCAAATTAAGTTAATCTTTAATCTCAAGGATTTAAGT 60
|||||
DB 94 ATGAAAAAACAGGTTTCCAAATTAAGTTAATCTTTAATCTCAAGGATTTAAGT 153
OY 61 AAAAAGTCAAAAGATTACTGTCAGTTAGTGGAGCTTTTAAATGATCTTGCTTG 120
|||||
DB 154 AAAAAGTCAAAAGATTACTGTCAGTTAGTGGAGCTTTTAAATGATCTTGCTTG 213
OY 121 GTAACTTCCAGTGTGGTGAAGACGTTTGTGTTAGTAGAATCCTGAGCCAAAC 180
|||||
DB 214 GTAACTTCCAGTGTGGTGAAGACGTTTGTGTTAGTAGAATCCTGAGCCAAAC 273
OY 181 GCATTAATCCAGATTCAAGTTGCGAATACAGATGATATGATATGATCTTAATGA 240
|||||
DB 274 GCAATTAATCCAGATTCAAGTTGCGAATACAGATGATATGATATGATCTTAATGA 333
OY 241 GGGCATCCATTTTAAACAGTTTAAAGTATAGATACACAGATTTAAAGGTTAACTTGAAGA 300
|||||
DB 334 GGGCATCCATTTTAAACAGTTTAAAGTATAGATACACAGATTTAAAGGTTAACTTGAAGA 393
OY 361 AGTAGTGTAAAGTGTATTAAGTAAACATGATGATGATCTTCAAAATTTGAAGATTAT 420
|||||
DB 454 AGTAGTGTAAAGTGTATTAAGTAAACATGATGATGATCTTCAAAATTTGAAGATTAT 513
OY 421 GCGATGAGCCCTAGAAATTAAGCGAGATGAGCTTAATGAGAAGTATGATATAT 480
|||||
DB 514 GCGATGAGCCCTAGAAATTAAGCGAGATGAGCTTAATGAGAAGTATGATATAT 573
OY 481 AATGACATCCACAAATGCCAATGATATTAAGGAGCTTGAACCCCTGATATGATC 540
|||||
DB 574 AATGACATCCACAAATGCCAATGATATTAAGGAGCTTGAACCCCTGATATGATC 633
OY 541 AGAGTTACACAGAGGGGGTATGCTATCTGATATGATATGCTATTTGATATCCAGAT 600
|||||
DB 634 AGAGTTACACAGAGGGGGTATGCTATCTGATATGATATGCTATTTGATATCCAGAT 693
OY 601 GAAAGTTTAAAGGAGTCAAGAAATGATGTTAGTATGTTAGTATCTTCAATTTATCTTGATG 660
|||||
DB 694 GAAAGTTTAAAGGAGTCAAGAAATGATGTTAGTATGTTAGTATCTTCAATTTATCTTGATG 753
OY 661 CGTCAAGCTTTGAAGCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
|||||
DB 754 CGTCAAGCTTTGAAGCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 813
OY 721 CCGGATGATTTTGAAGCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
|||||
DB 814 CCGGATGATTTTGAAGCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873
OY 781 GGAATCAAAATCTTTTGAAGTGTGTTAGTATCTTCAATTAACCAACCACTCCAGAGAC 840
|||||
DB 874 GGAATCAAAATCTTTTGAAGTGTGTTAGTATCTTCAATTAACCAACCACTCCAGAGAC 933
OY 841 CCACCAATGCTCTCCAAATCAACCTCAACGACCTTCACTACTTATTAAGAAATATGCTATA 900

[illegible]

gene
CDS
1872..3968
/gene="prf1.12"
1872..3968
/gene="prf1.12"
/note="fibronection-binding protein; protein F1; SfbI"
/codon_start=1
/transl_table=11
/product="protein F"
/protein_id="AAL6405.1"
/db_xref="GI:19224134"
/translation="MSSYMFARGEKMNMMFLNKEAGFLVHTKRRRAVTVGVF
LLACAGAIIGFQVAYADAEKTVNFKSPDPYWGDSYRGIFARYHNKVLKKS
KEYQAVFNLTKYRPRTYSTNNFKYKIDGSGSAFYANPVLDENDKLEKLL
NVIYNGKSNANGMNGIEDNALIVTQNAIWIYSDSAPLVNDYKMMERERKNEISE
SOVTLKREALKLIDPNLEAIAANKIPSGYRLNLFKSENEYQNLISAEEVDDPPK
GDTSEHNKTPPELDGPIIPDPKRPDSSSEPALPLPMLDELGEVEVPVPSLESPALP
PLMELDGEVEVPVPSLESPALPLPMLDELGEVEVPVPSLESPALPLMELDGEVE
VPEVPSLESPALPLMELDGEVEVPVPSLESPALPLMELDGEVEVPVPSLESPAL
YITEVYNGOONPVNDIKKLPLENGSFGNMVETEDTKRPVYLMGSGSESVETKDTQ
MSGOTTPOVETEDTKRPVYLMGSGSESVETKDTQMSGOTTPOVETEDTKRPVYLM
GGSGSESVETKDTQMSGOTTPOVETEDTKRPVYLMGSGSESVETKDTQMSGGS
ETVIVIEDTKRPVLFPHFDNNEPKVEENREKPKNTIPLPATGDIENVLAFLITLIS
VLSIFSLKKNKKV"
5326..7596
/gene="cpa.12"
5326..7596
/gene="cpa.12"
/note="putative collagen-binding protein"
/codon_start=1
/transl_table=11
/product="Cpa"
/protein_id="AAL6406.1"
/db_xref="GI:19224135"
/translation="MNNKLLQKODAPRVSNRKPQVTLVGVFLMFLVSSMRGA
QSIPEEKRIEVSVPKIKSPDAYPYGYDSYDSSHRYERFVADLRNKLNSKS
YQVCFNINSHYPRNKAQSPQWFKRVGDTGDFVNTAQTPIRGESLNNKLISIMN
AYPRNANGYMDKIEPLNALIVTQOAVVYSDSGNFKITLAEVPSLPAPQSPRP
REAYSKLISDDETSKKNLIPQSKLNIYVPOKSYONILISAEVPSLPAPQSPRP
PNOTKRSVIRKKAEGDYSKLEGATLRITGEDIIDPOEKVROSKTGKIEISNT
YTLFESSPDGKIAEPIKFRVNNKVFIVOKDSQVENNKVEAPYSVEASVSDWD
SNYINPEFTPYKFEYAKNKRDSOVYVCFNADLHSPSEDEGGTIDIDISTMKCY
KYHTAGSDLEFKALRPDNPEDLFHKIKVLEKGYNNKGDYSGNLTQOFRAQL
AIYFTSDTLKTLKTYNNGKYGHFESMDKELATVAKELINAQNSAPQITNLDF
VPMNSKQSLIGTEVHPDLDVDIRMEDKQEVYPTVHSLTVKTVYGEIGDRTKGRQ
PELEKDTGQPIVNTLKTNNODIVAKDQSRILKAGDTIRIEGLTGTSTYLTKE
AKDYIVTVNKRVSQEAQSASENVYADKEVIFENRKDLVPPITGDTGLWLLVP
FGLVWFGGRGLKND"
7589..8110
/gene="lepa"
7589..8110
/gene="lepa"
/note="putative signal peptidase I"
/codon_start=1
/transl_table=11
/product="lepa"
/protein_id="AAL6407.1"
/db_xref="GI:19224136"
/translation="MTNYLNRNLNENPLFKAFIRLVKLSITIGLGYTLFOYVFGVMIV
LTMQSPVASGQVLYRLTDRYHINDVVVYVEDNLLKGRILAAQAGDVSFTQEGG
LNTQSPPEKEVEPYLTPHSSGPNFKVPTGTGYFLINDYREERLDSRYGALPINOI
KGRISTLLVRG"
8132..9160
/gene="eflSL.A12"
8132..9160
/gene="eflSL.A12"
/codon_start=1
/transl_table=11
/product="eflSL.A"
/protein_id="AAL6408.1"
/db_xref="GI:19224137"

gene
CDS
9176..9901
/gene="eflSL.B"
9176..9901
/gene="eflSL.B"
/codon_start=1
/transl_table=11
/product="eflSL.B"
/protein_id="AAL6409.1"
/db_xref="GI:19224138"
/translation="MMFTIVQVINKAIDTLILFCLVYFLAGEGLMDSYHLVQADA
SNFKFKTAQOQPEFDLLAINEVIGMLNIPGTHIDYPLVQGTNLEYINKAVDGSV
AMSGSLDLDRNRNDEDDYSLYGHMAGNAGEIPKLRKDFEKHAKAIETKE
RKLTIVIFACLTADFDOLVFNPNAITNODQOQLVDYISKRSKQKPYKLKHHTF
VAPSTCFNFTDNRYIVGTIQE"
9936..10505
/note="orf2"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAL6410.1"
/db_xref="GI:19224139"
/translation="MFSVVMILMLAFNQTLYAKDSTVOTSISENVYERADSTPF
SIALESIDAMKTIETITAGSGKASEPLFTFTVGTQTVRYVQPSQNKYQADTVF
DVLVYVYDEDDGLVAKVISIRAGDEKSAITPKRLVPIPPROPNIPKPLPIAG
EVKSLGILSIVLGLVLVLYKRLKL"
complement(10664..11869)
/gene="msmRL"
complement(10664..11869)
/note="msmRL"
/note="AraC-type DNA-binding domain-containing protein"
/codon_start=1
/transl_table=11
/product="msmRL"
/protein_id="AAL6411.1"
/db_xref="GI:19224140"
/translation="MVIFDLKHQTLHSQDPISVNSQDRAILOVYGGNDYLLCYQ
FLKHALPQAQVIVFYEGLFEEGFMFLPCHYIIAIGPPYPSLKKDYOQLANFL
KHSRSEKLEKLSMALVPHPEPINNVNMLIADAFEDQETTCOOTIHLQSHSG
MTADPDIHLLKHSKASQDLPVLEHNLIMLQNLKGNOLLEOTNRPSSITSS
SISALREKULYIYILRLLEPSFVENTDVKRYSLYKYMALNEBSDLKYLIRIC
AAITHESESILFNKSIDSRQMYNSVLRHYDSHLKSLKYSQDAKRLVLSHLSRYFK
KYSNVSLOHYILSPKLEAQLLRGLPGEVAKSLYEDTTHFKIFKKYTGSISSKD
YLAKYRDNI"
12254..15739
/gene="prf2.12"
12254..15739
/gene="prf2.12"
/note="fibronection-binding protein; pfprI"
/codon_start=1
/transl_table=11
/product="protein f2"
/protein_id="AAL6412.1"
/db_xref="GI:19224141"
/translation="MTQKNSYKLSFLSLTGTGLGLLVFGLSGVSVGHAETRNGAN
KQGFELKRVKDNKPLPGATFSLTSKDGGKTSVQFTSNDKQIVAOINLOPQTYLK
ETAPAGDYKTSNNKTVYVYENGTYKLVENPYNGEILISKSGSDVSSLOLEPKKSY
VSKYKTEVSSGAADYRNAAVEKMSFKOKDKSETINPGDTFVLOIDRLNPKGI
SODIKTIYDSANSPALIGKYHAENQULITFDIYAGDKVQLSDELSTFLENKYL
ENTSISFKSTIGQETTYKGTNVNLIGNSTKESNVITNGLSNVGSGISTVTEGE
FWYVYVNPNRINIPIATMNLWFGARSNVTSOLENDANTSSALGEIOYVEPDEK
LPSYGVYKTLIRNDITAGLNGEQMTRORIDFGNNIONKAFIIRYVGTQDQSGK
PLVVOISLASFGCASEYAFTPVGWNVYRONEIALSPSKSGSGSEFTKPSITVNL
KRYAOLRBFKMSNDVPLPAAFEILSSNGNSQKLEASSNTOGEVHFKDLTSGTYLY
ETKAPKGYOQVETKLAIVYDTTKPAEBEYTWGSPSSSVKVEANKRVTYVNHKETITF
SGKKIENDRPDORPAKIOVOLLQNGQKMPNOIQETKNDKDSYHFKDLPKDAKQOE
YKSYVEEVNPDGYKVSYLGNIDIFNRETEFEVQNNFNLPEGNALIKQSGSKIDE

/note="extended promoter 2"
complement(6161..6166)
/gene="cpa"
/note="promoter 2"
6231..7898
/gene="nra"
6231..6241
/gene="nra"
/note="extended promoter 2"
6297..6302
/gene="nra"
/note="promoter 1"
6315..6325
/gene="nra"
/note="extended promoter 1"
6363..7898
/gene="nra"
/note="negative regulator of group A Streptococci"
/codon_start=1
/transl_table=1
/product="Nra"
/protein_id="AAC97147.1"
/db_xref="GI:4028948"
/translation="MPYVKKKDSFLVETYLEOSIRDKSELVILFFKSPITFSVAK
OTGTAVALKYCKEELDPFGNNIDITIKKGIICCVKPKREYILHQLDFTSILK
LVFFIKNGTSSOPLIKSKYFLSSSAYRLRESLILKREPGLRVSKNTIVGEYRI
RYLIAMLYSKFGIYIPLDHLDNOIYRFLSOSATNLTSPWLEEPSFYNNMLALSM
RKQFAVISIPQRIYFRLQKLFIDCLTRSSROYENASFJFSQGDLYFLYIT
NNSFASLOMTPOHIEFCCHIEKNDPRLLIEPLKRLPOLNHSKODLIKAMFSKS
FLFLQHFVIEIPESLPYTGNSLKKALKNIYOMLAPGKRLHLEKLOLFCSH
IEOLKHKOPALYVLISNFINAKLITDIPYRSDKGIFHSFYLLRDLIDYIPSL
KPOLYTHSRILIPVKNDLYKGVYVAEFSDKPDYIASIONLITQLKDKKYODFLNE
QLQ"
8006..8016
/note="ORF 5; extended -10 signal"
8024..8857
/note="ORF 5"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAC97154.1"
/db_xref="GI:4033717"
/translation="MDKIISIAOSGAFRAYVLDSTFVALAOKHNLTSSRYALGR
TLIANOLIANOKGDSKIYKVIKIDGSSFGHISADPKGHYKGIQNTQCDIKTERG
EVLGPRMGHRYTIDYGTGNPTISTPLTIGELIEDRAYITLSEQDPSALGLV
LIDENDVKVAGGFMOVLPBASEELIAREKRLQEMPALSHLSAHNDALLEAY
GDEPYKRLSEPLSFQDCSREREALMTLPKADLOAMIDEDKGAIEVCOFCGCTLP
1"
8884..9861
/note="ORF 6; similar to Acospirillum brasiliense nifH"
gene"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAC97155.1"
/db_xref="GI:4033718"
/translation="MILNSSFRIQDVEIDHRYVLAPMAGVTSAPRTIAKEGAGIV
VMEIISERGLYNNKTKLHMLHIDENHPISIDLFGDAGLKAAPFIQNTADIV
DINMGCPVNRVYVNEGAKVLRDPDKIYHLYKVEYTSVLDIPYKMTGADSSLAIE
NALAESAGVSALAMGRTNOMYGTCDHEHTLAKATIKPIEFINGDVRSYODAK
FMIEITGVADVAMGRAMNNPYLFTOINHFETGSELDPDPEAKKLDIADHLKRLIN
LKGETTAVREFRGLAPHYLNGTAGAAKRAVSAETLAELAEYFEYR"
9881..10522
/note="similar to Lactobacillus acidophilus da/dg-kinase"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAC97156.1"
/db_xref="GI:4033719"
/translation="MLIVLAGTIGAGSSLAALGELHGTIDVEYEAVDNNPVIDFYIO
DPKRYAFLQIYFLNKRKSIKEAYQADNNILDSISIEDELFLANKNNVKTIEDI
IYELLANMLEELEGMPKRRPDLIYIDVFCKMLERIEKRGSEFQVDNPSLEQY
HOVHGEPYTWYEDYKVSPPMKIDONSIDPQNPQDPLATVTKMIDTKIKELHLL"

repeat_region 10598..10681
/note="putative transcription terminator"
/rpt_type="inverted
complement(10782..>10826)
/gene="ssba"
Query Match 31.7%; Score 720.8; DB 1; Length 10826;
Best local similarity 61.6%; Pred. No. 8.7e-113;
Matches 1372; Conservative 0; Mismatches 787; Indels 69; Gaps 11;
gene
61 AAAAAGCTCAAAACGATTTACTGTCATTTAGTGGAGCTTTTAAATGATTCGCTTG 120
11
5894 AACAAAGAGACCAACAGACGATCGATGATGAAAGATTTTGGACCTTGACCTG 5835
121 GTAACTTCATGTTGGTCTTAAGACCTTTTGGTTTGTAGTAATCCGACCCAAAC 180
11
5834 ATGAGAAATAGTAGGTTTCTATCAGACCTTCGAGCTGAAGAAATCAGTACAAA- 5776
181 GCATTAATCCAGATTCAGTTCGGAATACAGATGATGATGATGATGATGATGATG 240
11
5775 -----TAGACAAAGCTCAATTCAAGATTATCCGTGATAGGCTATGATTCCTAA 5721
241 GGCATCCATATTTTAAACAGTTTATAGTACACACAGATTAAAGGTTAACTTGAAGA 300
11
5720 GCGTACCACGACTATAGTCCGTTAAAGACTTACCATATTTAAAGTAAATTTAGAGGA 5661
301 AGTGAAGTTTCAAGTTTATGCTTAAATTTAAAGAAAGCATTTCTCGATCCAGAT 360
11
5660 AGTAAAGATTTATCAAGCTACTGCTTATTTAAACAAACATTTCCATCCAGTCAGAT 5601
361 AGTACTGTTAAAGTGTATATAAAACATGATGATCTCTACAAAATTTGAAGATTAT 420
11
5600 AGTGTAGATCAGATGATATAAAACCTTAAAGCACTAATGAACCTTATATCAAGTTA 5541
421 GCGATGAGCCCTAAATATACGAGATGAGCTAATACAGATTAACGAGTGTATGATAT 480
11
5540 GCATATAACCAAGATATGAGAGGAGGAGCAGTTACAAACAAATATATGAGATTCTCAT 5481
481 AATGACATCCACAAATGCCAATGATATATGAGAGGCTTGAACCTTGAATGATATC 540
11
5480 AATGATATCTTAATTAATCGTAATGAGTATGAAAGGATAGATCCCTTAACGCTAT 5421
541 AGATTTCAACAAGGCGGCTATGATACATTCGATATATGCTATTTCTTAATCCAGAT 600
11
5420 TTAGTACTCAAAATGCTATTGTGTA-----TACTGATTTAGCTCAAAATTAACCGGAT 5367
601 GAAAGTTTAAAGAGGAGTCAAGAAATGATGTTAGTACTTCAATTTCTTTGATG 660
11
5366 GAAAGTTTAAAGAGGAGTCAAGAAATGATGTTAGTACTTCAATTTCTTTGATG 5307
5306 CGAAAGCTTTAAAGAGGAGTCAAGAAATGATGTTAGTACTTCAATTTCTTTGATG 5247
721 CCGGATGATTTTCAAGTATGATTTTCAAGTATGATGATGATGATGATGATGATG 780
11
5246 CCATCAGGTTATGCGTTAAAGTATGTAATGTAATGTAATGTAATGTAATGTAATG 5202
781 GGATACCAAAATCTTTGAGTGGTGGTATGATGATGATGATGATGATGATGATGATG 840
11
5201 ---TTCCAAATCTTTGAGTGGTGGTATGATGATGATGATGATGATGATGATGATG 5146
841 CCACCAATGCTTCAAAATCAACGCTCAACGCTTCAAGTATGATGATGATGATGATG 900
11
5145 -----AGAGCTCCGCGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 5091
901 GGTGATTAATCTTAATGCTTGAAGTCAACATGATGATGATGATGATGATGATGATG 960
11
5090 GGTGATTAATCTTAATGCTTGAAGTCAACATGATGATGATGATGATGATGATGATG 5034
961 AGTTTCAAGGAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
11
5033 GGTTCACAGAAAGAGCTTCAAGATATGATGATGATGATGATGATGATGATGATGAT 4974

OY	1021	GGAACTTATCTTTAACTGCAATTCGATCTCCAGCTGGTTATGATATCGCAGAGCCAACTC	1086
Db	4973	GGGACTTATACCTTTACACGAAACATCACTCAGATGGATTAATAATTGCGGAGCCGAT	4914
OY	1081	ACTTTTAAGTTGAAGCTGGCCAAAGTGATACAT - - - TATGATGGAAGAAACAGATGTAA	1137
Db	4913	AAGTTTAGAGTAGAGAAATTAATAAAGTATTTATTCGTCCAAAAGATGGTTCTCCAAAGTGA	4854
OY	1138	AATCCCAATTAAGAGATAGTAGAGCCCTTACTCAGTAGAGACATTAATGATTTT - - - GAA	1194
Db	4853	AATCCCAATTAAGAGATAGTAGAGCCCTTACTCAGTAGAGACATTAATGATTTATGATGAT	4794
OY	1195	GAAATTACCGTTTAACTACACAAAACATATGCAAAATTTTATATGCAAAAAATTAATAAT	1254
Db	4793	GAAAGATACCTCTCGGGTTTACTTCCACAGGAAATTTCTATTACGTACAAATAAAGAT	4734
OY	1255	GGAAGTTCACAGGTTGTCTATTGGCTTTATGCTCAGATCTTAATAATCCCAAGACCTGTGA	1314
Db	4733	AAAAGTTCACAGGTTGTCTACTGCTCAATGCTGATTTTACCTCACCCAGCTGACTCATAT	4674
OY	1315	GATGGTGGGAAACAAATGACTCCAGA - - - CTTTACAAACAGAGAGATTAATAATACATCAT	1371
Db	4673	GATAGTGGTAGACATTAATCCAGATACATAGTAGAGTAAGAAAGATCAAGTACACACAT	4614
OY	1372	ATTGCAAGTCTGTACCTCTTTAAATATATCTGTGAACCAACAGATACCGATCTGCACACT	1431
Db	4613	ACGGCAAGTATGACTGTGTTTAAATATCCGCTAAGACCGAAGATATCAAAATCCGAAACAC	4554
OY	1432	TTCTTAAACATATCAAAAAAGTATTTAGAGAAGGGTTACAGGAAAAAGACAAAGCTATT	1491
Db	4553	TTCTTAAAGCAATTAATAAAGTATTTAATAAAGGCTTCAAGAAAAAG - - - - - TGAT	4500
OY	1492	GAGTATAGTGGTCTPACTAGACACAAATTCGTCGGCTACTCAGTTAGCAATATATTAT	1551
Db	4499	AGCTATATATGATTAAACAGAAACACAGTTTCGGCGGCTACTCAGCTGTATCTATTAT	4440
OY	1552	TTCACTGATTAAGCTGATTTAGA - - - - - TAAAGTAAACTTAAAGACTATCAT	1599
Db	4439	TTTACAGACATCGACTTTAAAAACCTTAAAACTTATAACATGGGAAAGGTTACCAT	4380
OY	1600	GGTTTGGAGACATGATAGTATGATCTTACAGTTCGTAAATCCCTGTGAATATGCT	1659
Db	4379	GGATTTGATCTATGATGATGAAAAAACCTTACTGTACAAAAAGATTTATTTACTTATGCT	4320
OY	1650	CAAGATAGTAATCCTCCACAGCTTAAGTACCTGATTTTCTTATATCCGAATTAACATTA	1719
Db	4319	CAAAATGCGAGCGCCCTCAACTTAACAATCTTGATTTCTTGATACCTAATTAATAGCAA	4260
OY	1720	TATCAATCTCTTATTTGAACTCAGTGGCATCCAGAGATTTAGTTGATTTATTCGTATG	1779
Db	4259	GACCAATCTCTTATTTGGGACGAATAGCCATCAAGTATGTTGGTTGAGCTATGCTATG	4200
OY	1780	GAAAGTAAAA - - - AAGAAATTAATACGTATCATATTTAATTAACATTGAGAAAAAGGTG	1836
Db	4139	GAAATTAATAAGCAAGAGATTTATTCAGTACTACAGTTTGACAGTGAATAAABACGTA	4140
OY	1837	ACTGGTTTAACTGGTGACAGAACTTAAGATTTTCATTTTGAATTTGAATTTAAAAATTAAT	1896
Db	4139	GTCGGTAGTGGGAGATTAATAACCTTAAGCGTCCCAATTTGAACCTTGAGTTTAAAGATTA	4080
OY	1897	AAGCAAGATTTGCTTTCTCAAACTGTTTAAACAGATTAACAAACCTTCGAATTTAAAGAT	1956
Db	4079	ACTGAGACGCTATTGTTTAAACACTCTTAATACTAATATTAACAAGATTTAGTAGTAAAGAT	4020
OY	1957	GGTAAAGCAACCATTAATTTAAACATATGGGAAAGTTTACACCTTCAAGTTTACCGAGA	2016
Db	4019	GGGAAATATTTATTTAATCTTAACACATGCTGACACCATTAATAATTAAGAAAGTTTACCGACG	3960
OY	2017	GGTATATCTTACCTTGCCAAAGAAACAGATTCTGAAGCTTATAAGGTTTAAAGTAAATAGC	2076
Db	3959	GGATATTTCTTATCTGAAAGAGCGCTGAAGCTTAAGATTTATATAGTATACCGTTGATATAC	3900

QY	2077	CAGAAGTAGCAAAATGCTACAGSTTTCCAAAACAGGAATACACAGATGACACACTTCT	2136
Db	3899	AAAGTTAGTCAAGAAAGCCAGCTACGTAGTGAAGGTATACAGAAAGACAAAAAGTCACT	3840
QY	2137	TTTGAAAAATATTAAGAGAGCTGTTGTTCCATCAGAGAGATTGATCCAAAAGATCAATGGCTAT	2196
Db	3839	TTTGAAAAACGAAAAGATCTTGTCGCCACCACTGGTTTGACACAGATGGGGCTATCAT	3780
QY	2197	CTACGTTTGGATTTGATTCGCTGTGATATCAGTTGGGAGATCTGGGGAATTCACACGATAAAG	2256
Db	3779	CTTTGGTTGTTATTACTCTGTTCCACTTGCTGTTATATGCTGTTGGCTATTTGGTGTGTAAGGG	3720
QY	2257	ATAAGAAA	2264
Db	3719	TTAAAAAA	3712
RESULT 5			
AE014138			
LOCUS	AE014138	52900 bp	DNA
DEFINITION	AE014138	AE014074	linear BCT 19-JUL-2002
ACCESSION	AE014138	AE014074	
VERSION	AE014138.1	GI:21903795	
KEYWORDS			
SOURCE			
ORGANISM	Streptococcus pyogenes MGAS315		
REFERENCE	Streptococcus pyogenes MGAS315		
AUTHORS	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.		
	1 (bases 1 to 52900)		
	Beres,S.B., Sylva,G.L., Barblian,K.D., Lei,B., Hoff,J.S., Mammarella,N.D., Liu,M.-Y., Smoot,J.C., Porcella,S.F., Parkins,L.D., McCormick,J.K., Leung,D.Y.M., Schlievert,P.M. and Musser,J.M.		
TITLE	Genome sequence of a serotype M3 strain of group A Streptococcus: Phase-encoded toxins, the high-virulence phenotype, and clone emergence		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99 (15),	10078-10083 (2002)
REFERENCE	2 (bases 1 to 52900)		
AUTHORS	Beres,S.B., Sylva,G.L., Barblian,K.D., Lei,B., Hoff,J.S., Mammarella,N.D., Liu,M.-Y., Smoot,J.C., Porcella,S.F., Parkins,L.D., McCormick,J.K., Leung,D.Y.M., Schlievert,P.M. and Musser,J.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-JUN-2002)		
	Rocky Mountain Laboratories, NIAID, NIH, 903 South Fourth St., Hamilton, MT 59840, USA		
FEATURES	Location/Qualifiers		
source	1..52900		
	/organism="Streptococcus pyogenes MGAS315"		
	/mol_type="genomic DNA"		
	/strain="MGAS315"		
	/serotype="M3"		
	/db_xref="taxon:198466"		
	/note="group: A"		
	77..1015		
gene	/gene="comYA"		
	/note="synonym: SpvM3_0078"		
	77..1015		
CDS	/note="best non-GAS blastp hit: gp AAC45310.1 (U081957) putative ABC transporter subunit ComYA [Streptococcus gondii]"		
	/codon_start=1		
	/transl_table=1		
	/product="putative ABC transporter subunit ComYA"		
	/protein_id="AA078685.1"		
	/db_xref="GI:21903796"		
	/translation="MVAALAKAIIAAKEOVAODIYLIPRADOYDLFLRIGDERRLVDVYSDRAAPLISHPKPVAGMIVGEKRRQVGSQDKLSKDKQLSLRSVGDYGGSSIVRIILRHONKSVHYWFDGLTKVANGVGGILYFAGPVGSGKTTLMYOLISSYHQAQVDSIEPVEIKHQIILQLOVNDIDIGMTYDNLILKLSLRHPDILVIGEIRDSQIARAVIRLSLTPAMFVSTVHAKSISGVYARLLELGTVAEILSNCAILAYORLNGGALIDSTV		

gene
951. .1985
/gene="comYB"
/note="synonym: spym3_0079"
951. .1985
/gene="comYB"
/note="best non-GAS blastp hit: gb|AAC45312.1| (U81957)
putative ABC transporter subunit ComYB [Streptococcus
gordonii]"
/codon_start=1
/transl_table=11
/product="putative ABC transporter subunit ComYB"
/protein_id="AAM78686.1"
/db_xref="GI:21903797"
/translation="MISFLRODISIPSKLSKRSKHQHFQIOLANLSTGSEFAE
VIAIQRNLNOMARIRKRTVEYITPILLLFLPYMAGLRVLYPQLETOHNTFTLN
HFAFTGFCGSLILLEGVWLMRWSRSLKLSRLRYPLGKLLQYLTYSAREW
GTLIGGLDMLTILDMAIEKSLMLKELADIRNSLEGQAFHIVKATYPPKFKLSL
MIYGEGLKIGALELEIYAQSEWEPFSQLYQVYQLIQAIFLVAVATVIMYAILL
PIYNNNGGIF"
1987. .2313
/gene="comYC"
/note="synonym: spym3_0080"
1987. .2313
/gene="comYC"
/note="best non-GAS blastp hit: gb|AAC45312.1| (U81957)
comYC [Streptococcus gordonii]"
/codon_start=1
/transl_table=11
/product="putative competence protein"
/protein_id="AAM78687.1"
/db_xref="GI:21903798"
/translation="MINQNNLRHKKLKGFTLEMLIYIIVSYIMLLFVNLSTOKD
RYTERGNAVYKLVENQAELEYLSQSKRPSLSOLKADGSTTEKQEKAYQDYDKHNE
KALNSN"
2273. .2716
/gene="spym3_0081"
2273. .2716
/gene="spym3_0081"
/note="best non-GAS blastp hit: gb|AAL00663.1| (AE008550)
Competence protein [Streptococcus pneumoniae R6]"
/codon_start=1
/transl_table=11
/product="putative competence protein"
/protein_id="AAM78688.1"
/db_xref="GI:21903799"
/translation="MTNKKMKRPVLAIKAFTLLETLLSYMSFIIILGSVPVTKSYQ
KVEEHFFSFHEHLRHQOKLAILQOKROYLDISSTKIVTEGNSLTVPKSTIVNHPYR
LVLDQGNHSLAKIIFDMTRRKYQFYIGSGNYKTSQSLHSP"
2673. .2957
/gene="spym3_0082"
2673. .2957
/gene="spym3_0082"
/note="best non-GAS blastp hit: gb|AAK06217.1| (AE006440.3
(AE006440) hypothetical protein [Lactococcus lactis subsp.
lactis]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAM78689.1"
/db_xref="GI:21903800"
/translation="MGIIRKQVAYIALESIMANGILFSYIIVLSSLOSOALITYY
RKQEKINLALMAVQTRTKEMTLNGCHITTLIRDRYISIHDBEGEMKTE"
2950. .3384
/gene="comYD"
/note="synonym: spym3_0083"
2950. .3384
/gene="comYD"
/note="best non-GAS blastp hit: emb|CA75315.1| (Y15043)
homology to ComYD from Streptococcus gordonii, and ComGD
from Bacillus subtilis [Lactococcus lactis subsp.
cremoris]"

gene
3368. .3694
/gene="spym3_0084"
3368. .3694
/gene="spym3_0084"
/note="best blastp hit: gb|AAK33225.1| (AE006481)
hypothetical protein [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAM78691.1"
/db_xref="GI:21903802"
/translation="MPEFKNLKAGILLIYAMFNASIFLLVQVLYLSQVLTALHREXOATF
DYVKARLIEIYVQDHRKASNPVFFKGGVYICREKREKMLLVKIDQOQYQFEYIK
PROXSQ"
3792. .4745
/gene="spym3_0085"
3792. .4745
/gene="spym3_0085"
/note="best non-GAS blastp hit: gb|AAK06112.1| (AE006431.4
(AE006431) hypothetical protein [Lactococcus lactis subsp.
lactis]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAM78692.1"
/db_xref="GI:21903803"
/translation="MTFEKITEAVYQLLENQCLIENTDKTHIYDAIVEONSPFLGAEQ
ASPQVQNSDKLALCLTRKEEMKRAYQFLIKAAQEQLOQANHQFPDAIGFLLYLL
EQLSDKDSLEVLIEGSGTGNLAQTLNNTSKSDYQIEDDLIIIDIASIAELMDS
AHFIQEDAVRPOLEKESDIYISDLPVGYYPNDIYAKVYASDKRHYAHHLMEOSL
KYLKQGFAPILAPVNLITSPOSOLIKOMLQDAOVYALIMLPDSIFGRHSNKSIIY
LOKQIDPMETFEYPIRDLKLAENIHDFENFKKMLSNV"
4804. .6000
/gene="ackA"
4804. .6000
/gene="ackA"
/note="synonym: spym3_0086"
4804. .6000
/gene="ackA"
/note="best blastp hit: gb|AAK33227.1| (AE006481) acetate
kinase [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="acetate kinase"
/protein_id="AAM78693.1"
/db_xref="GI:21903804"
/translation="MSKTIAINAGSSSLKMWQLYQMBEAVLAQGIITERGLKDSISTV
KYDGKKEQIILDIHDTFAVKILINSSLLKWLQYQMBEAVLAQGIITERGLKDSISTV
VNDKYLEQIEELSVLALPNPQAAAGIRFROLDLPITSQVCDTSPHTSMARHTYLY
PIPOKYVYTDYKVRKGAHGTSHKYYVNOEAKMGRPLEIKLITAHNGVSTANVH
GKSYDTSMKFTPLACGPMKSTRSDIDPAIIPYIIEQDPKADADVNNMLNKSGLSG
VSGISDMRDIEAGLOEDNPDAVLANNITFDRIKCTIGQFAVLNGADALVPTAGME
NAPLMKQVYIGGLTWFGMDIDPEKRVNGYFISTPESKRVYVISTDELCIARDE
RLKNYK"
6186. .6494
/gene="spym3_0087"
6186. .6494
/gene="spym3_0087"
/note="best blastp hit: gb|AAK33228.1| (AE006481)
hypothetical protein [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAM78694.1"
/db_xref="GI:21903805"
/translation="MYVFIEYNDLKDYLITVYDERNRAIRKOTQVCFYLMIVLVAL
FEFFPLIRDLVSYKYLPIEFYVYIILTSIVAKGELGKITNTSKQPPEDTSOYVC
FEFFPLIRDLVSYKYLPIEFYVYIILTSIVAKGELGKITNTSKQPPEDTSOYVC

Db 17667 TTTAAACCGAAAGACTGTGCTCCACCACTCG-----TTTGACACAGATGGGCGTAT 17722

Oy 2197 CTAGCTTGATGATGATTCCTGCTGATACGATTGGGGAT 2234

Db 17723 CTATCTTGTGTTGTTATTAATTACTGTTCCATTTGGGTTAT 17760

RESULT 6
AP005141 311600 bp DNA linear BCT 13-JUN-2003
LOCUS Streptococcus pyogenes SSI-1 DNA, complete genome, section 1/6.
DEFINITION AP005141 BA000034
ACCESSION AP005141.1 GI:28810157
VERSION
KEYWORDS
SOURCE Streptococcus pyogenes SSI-1
ORGANISM Streptococcus pyogenes SSI-1
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.

REFERENCE 1
AUTHORS Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y.,
Okahashi, N., Kawabata, S., Yamazaki, K., Shida, T., Yasunaga, T.,
Hayashi, H., Hattori, M. and Hamada, S.
TITLE Genome Sequence of an M3 Strain of Streptococcus pyogenes Reveals a
Large-Scale Genomic Rearrangement in Invasive Strains and New
Insights into Phage Evolution
JOURNAL Genome Res. 13 (6), 1042-1055 (2003)
MEDLINE 22683278
REFERENCES 2 (bases 1 to 311600)
AUTHORS Yamashita, A., Nakagawa, I., Kurokawa, K., Nakata, M., Tomiyasu, Y.,
Yamazaki, K., Okahashi, N., Kawabata, S., Yasunaga, T., Hattori, M.,
Hayashi, H. and Hamada, S.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2002) Ken Kurokawa, Osaka University, Genome
Information Research Center, 3-1, Yamadaoka, Suita, Osaka 565-0871,
Japan (E-mail: ken@gen-info.osaka-u.ac.jp,
URL: http://www.gen-info.osaka-u.ac.jp/, Tel: 81-6-6879-8365,
Fax: 81-6-6879-2047)
COMMENT genome project
This clone was isolated from a patient presenting with toxic shock
like syndrome.

FEATURES
source
1. 311600
/organism="Streptococcus pyogenes SSI-1"
/mol_type="genomic DNA"
/strain="SSI-1"
/db_xref="taxon:193567"
232. 1587
/gene="SPs0001"
232. 1587
/gene="SPs0001"
/note="similar to GB:AAE71535.1 (AF255728) percent
identity 100 in 451 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="chromosomal initiator protein"
/protein_id="BAC63096.1"
/db_xref="GI:28810158"
/translation="MTENEQIFMNRVLELAQSOLKQATYEFVHADARLLKVKHIAIT
YLDQKELFWEKNLKDVLITAGFEVNAQISVDVFEEDLITEONQIRKINPKPOAL
NSLPTVSDNSKYSFENFIOGDENRNVAASIAVANTPGTTYNDLPFWGSGIGKTH
LNAIGSVLENNRNARIKYITTAENFINFEVYIHRDPMDELKRFKRNLDLLIDDDIO
SLAKRTISGTOEEFFNFNHNKQIVLSDTDPHLNDELRLVTRKRWGLTVNI
TPPDFETRVAILNKIOENYTFPQDITEYIAGQGSNVRLGALKDISLVANFKOI
DTITVDIAEAIARAKQDPMIVIPIEIOAQGKGYVAKIKATKRONTVLAK
QVAFILAREMTDNLSPKIGKEFGGRDSTVLAHVAKIKNMISQESRIETETIKNKI
K"
1742. 2878
/gene="SPs0002"
1742. 2878
/gene="SPs0002"
/note="similar to GB:AAL96838.1 (AE009953) percent

identity 99 in 378 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="putative DNA polymerase III beta subunit"
/protein_id="BAC63097.1"
/db_xref="GI:28810159"
/translation="MIOFSIRTLTFIRNLRNATKRAISKNAIPILSSKIEVPTGYT
LTGNSGOISTENITPVSNNENAGLITSPGALILBASFFINITSSLPDININVEIEOH
QVLTLSGKSEITLKDVDOYPRLOEVSTENPLIKTKLKSIIAETAFASLOESRP
ILTCVHVILSNHKKDFKAVAFDSHMSQRLITLDNTSADFVYIPGKSLREBSAVFTD
IEIVFEFSPQILFRSEHISFYTRLEGNYPDRDLMTFEFEVEVFNOSLHAE
RAFLINATONGTIVKLTITONHISAHNNSPVGKVNEDLDIVSOSGSDLTISFPYTL
IESLKAISKSEVYKIHFLSPVAPFTLITFDEDEESTIQLITTPKRN"
2953. 3150
/gene="SPs0003"
2953. 3150
/gene="SPs0003"
/note="similar to GB:AAK33148.1 (AE006472) percent
identity 100 in 65 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="BAC63098.1"
/db_xref="GI:28810160"
/translation="MYOIGSFVEWKKRHHACVYKETGKKANQWVLRVAGADIKIOCTNC
OHVIMSRIDYDFERKRVLP"
3480. 4595
/gene="SPs0004"
3480. 4595
/gene="SPs0004"
/note="similar to GB:AAK33149.1 (AE006472) percent
identity 98 in 371 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="putative GMP-binding protein"
/protein_id="BAC63099.1"
/db_xref="GI:28810161"
/translation="MALTAGIVGLPNYKSTLFNATRKAGAEANYPATIDPNVGY
EVPDERLOKTELITPKKTYPTPEFDIAGIVGASGEGLGKPLANIREVAIYH
VYAFEDDENVMREGRDPAVDIIDLINLELILADLESINKRYARVEKMAFQD
KESVAEFENYLOKIKPVLDDGKSARTIETDEAKVVGRLITTKRPVLYVANVEDDY
ANPGIDYKQIRDPATENAEVYVISAAREEILLELDEDEKEFLAIGITGECYDK
LTRAHYHLGIVTFAQEKVRAMTRRGIKAPQAGIITHSDERGITRAVMSYDD
LMTYSSEKAVKEAGRLRECKEYVVOGDIMEFRNV"
4665. 5234
/gene="SPs0005"
4665. 5234
/gene="SPs0005"
/note="similar to GB:AAK33151.1 (AE006472) percent
identity 99 in 189 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="putative peptidyl-trna hydrolase"
/protein_id="BAC63100.1"
/db_xref="GI:28810162"
/translation="MYKMIYGLNPGSKYKTKNIGFMAIDNTIVKMLDYFTDKNF
KAQIGSFINIEKCEVAPPTFPNNSGIAVALLYINIDITDLIVYDDIDMEVSKL
RLRSGSGGNGIKSIIAHIGQOEFNRIRKVGIGRPLKGMVIVHWVGQFTEENIAI
SLITDRYVNAKFTLOENDPEKTKQKENG"
5237. 8740
/gene="SPs0006"
5237. 8740
/gene="SPs0006"
/note="similar to GB:AAK33151.1 (AE006472) percent
identity 99 in 1167 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=1
/product="putative transcription-repair coupling factor"

/protein_d="BAC63101.1"
/db_xref="GI:28810163"
/translation="MDILELFSQNKVKVQSMHSGLTLLIGRLVNLGSGSSKALAIASAY
LDDQKRIIVYTSQGVNEKTLASDLISLDEELVFCFADVAAEFTFASMDKALSRI
ETLQELRNPQSGOVLIVSGSRITLLPNDFVFKSOIOLTVGEDYSDTLTKLMTL
YKQSVQSLSPGSESRGDIOLVETLOELPYLRELFQSSYEDLVYAKNGKNGKIDIRK
GVFTNPASDLITFEASDFQGIQLEKALOTADDDKSYLEDLVYAKNGKNGKIDIRK
GSLFEKESLSDLYIPKGPPIFFDFQKLVKNARPLEIANLLTDLQKALSNIN
FYADYRELIRKPAFTFENFKGLNIFEDMHQLTOYAMQEFNFQPLILDEIKRY
QKNQTVYQVNSQVAYELKEKSPDYCRPLDVSANQIVSRESQIVIGASISGFYA
DEKLILTHEYELHYKKIRKARRSNISMAERLKDYNLEAGDVVANNHNGIGFLIE
ITLOIGHRDVTLYOXSNDISLPIQDQISLSIKYASADGPEKIKLNDGRQRTKQ
KVARVEDIADDLILYIAPERSQKGFSESDDDLQAFDDQAFVETEDQSLIKET
ADMESMOPDRLLVLDVVGKTEVAMRAAKFKAVNDKQYAVLPPTVLAQOHNENKA
RFPNPEVDVYLSRPSRKSQEQATELERYKGGIDILIGHRLISKVYVSDLSLIDY
LEORNGVKKETLKEITKEDVTDLTATAPILPTLHSMGLIRLSTIEMPTPTAPQ
TYLENNGVLVAYELIIRREDRGGQIFVYVYKQDTLEKKYAELOELVPAISIGYVHM
SEIQENTLIDPINDGYLVATTIITETVDISNVATLITENADHNGSLTLYDRGV
KSNINIAVAVYRQDLVETSEKLEAIKGETELSGEFKIAMRDLSTIGACNLTGA
SOSQIDISGEMFYSQLELOALASQKGLTKVARKQGTETNLQIDALPDYIADDERK
IDYKRIEISQREYLLNODLIDRFGEYPPQAVLLETALLKHNDNAFALVERK
NNQIVAFREVTLSFLNODLYEALSKTHLKAKKISHQCKLIDVFPVRQKQVRIILE
LMLFEERISEIKIRKNSYFK"
8878..9174
/gene="SPS0007"
8878..9174
/gene="SPS0007"
/note="similar to GB:AAK3152.1 (AE006472) percent
identity 100 in 98 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC63102.1"
/db_xref="GI:28810164"
/translation="MKRIYVETLDRKYLKVSRLIKRSGVAKEVADKRIYKGIILAK
STNKLKNDLIDHISFCNKLITFVRIEIKSTKEDALKWEIISETRITLNEBA"
9161..9532
/gene="SPS0008"
9161..9532
/gene="SPS0008"
/note="similar to GB:AAI96843.1 (AE009953) percent
identity 100 in 123 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative cell division protein (DivC)"
/protein_id="BAC63103.1"
/db_xref="GI:28810165"
/translation="MKKRSIYOLNNHYIKENLKKFEEESOKRRNFMGMLIVSMF
LEIPLTVLVKSYVDFEKNQOQVVLKRLKRYNELSKTKREKQALERLKDQDNIPYKAKR
AKYYSRSGEMTYPPLGLPK"
9529..9654
/gene="SPS0009"
9529..9654
/gene="SPS0009"
/note="unknown"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC63104.1"
/db_xref="GI:28810166"
/translation="MMMNIVKIEIAFLAFSDRKLAELOQENQKYEESQYIOKN"

Query Match	29.2%;	Score 664.4;	DB 1;	Length 311600;
Best Local Similarity	60.0%;	Pred. No. 1.3e-103;		
Matches 1318;	Conservative	0;	Mismatches 816;	Indels 64;
				Gaps 10;

QY 61 AAAAAGCTCAAAAGCATTTTACTGTCTACTTTAGTGGAGCTCTTTTAAATGACTCTCGCTTGG 120
111784 AACCAACGCGACGACAAACGCGATGTGATTTACTGGAAGTAAATTTTTCAGCTTTGTACTCTCG 111843
Db

QY	121	GTACTTCGATGGTTGGTGGCTAAGACTGTTTTGGTTTGGTTAGTGAATCCCGACGCCAAAC	180
Db	111844	ATAGAAATAGTAGGGGTTTTCTATCGAGGGGTTGGGCGCTGAAGAAACAGTGGCGAAA-	1119020
QY	181	GCATAAATCCCGATTCAAGTTCGGAAATTCACATGGTATGAGATATCAATCTATGTPAAG	240
Db	111903	-----TAAACAAAGCTCAGCTTCAAGATTATCCGTGGTATGCGCTATGATTTCTTTTAA	111957
QY	241	GGGCATCCATATTTATTAACAGTTTAGATAGACACACAGATTTAAGGGTTAACTAGAAGGA	300
Db	111958	GGCTACCCAGACTATATAGTCGGTTAAAGACTTACCATATTTTAAAGTAAATTTAGATGGA	112017
QY	301	AGTACAAGTATCAAGCTTTATTCCTTTAATTTAAGAAAGCACTTTCCTCGATACAGT	360
Db	112018	AGTAAAGGATATCAAGCATACCTCTTTAATTTTAAACAAAACCTTTCCATCCAACTGAGAT	112077
QY	361	AGTACTGTAAAAAGGTGATPAAAAAACATGATGGATCTGCACAAAATTTGAAGATTAT	420
Db	112078	AGTGTTAGATCCCAATGGATATPAAAAAAGTTGAAGGAACTATATGAANAACCTTTATCAAGTAA	112137
QY	421	GGCATGAGCCCTCAGATATTACGGGAGATGAGCTAAATCAGAACTTACGAGCTTTATGATAT	480
Db	112138	GCAGATTAACCAAGATATGAAAGACGAGCAGTTACAAACAATAATATTGAGATTCCTCAT	112197
QY	481	AATGACATCCACAAAATGCCAATGGTATTTATGAAAGGCTTGGAAACCTTGATGCTATC	540
Db	112198	AATGATATTCCTTAATGATCGTAAATGGATTAATGAAGGATATGATTCCTTAAACGCTAT	112257
QY	541	AGAGTTACACAAAGCGCGTATGATGATATCTCGATTAATGCTCCATTTCTATCCAGAT	600
Db	112258	TTAGGACTCAGAAAAGCTATTTGGTATGATCTGATGATGATTCCTATATTTCTATCTACTAGT	112317
QY	601	GAAATTTTAAAAAGGAGATCAGAAAGTAACTGGTATGACTTCTCAATTTATCTTTGATG	660
Db	112318	AAAGCTTTTCAACAGAAAGAAACACATCTTAATATAGATTCACAGCACTTACACATGAAG	112377
QY	661	CGTCAGCTTTGAAACCAACTGATGATCCGAAATTTGGCACTPAAATGCCAACAACAATT	720
Db	112378	CGTATGCGCTPAAAAAGATTAATTAATTCACAAAAGAGGTGAAAGAGTTTACCGCATCAGGTT	112437
QY	721	CCGGATGATTTTCCACCTAAGTATTTTGGTCTGAGGACACAGGAGATTAATATATATAA	780
Db	112438	CCCCGAATTTATCACTTAAGTATTTTTCATTT-----TCTGATAG	112479
QY	781	GGATACCAAAATCTTTTGGAGTGGTGTAGTTCTCTACTPAAACCAACAATCCAGAGAGAC	840
Db	112480	ACTTTCCAAAATCTTTTAAGTCTGAGATGATGTTCCGAGATACCTCCCAAAACCAAGAGA-	112538
QY	841	CCACCAATGCCCTCCAAATCAACCTCAACAGCTTAGTACTTATTTAGAAAGTATGCTATA	900
Db	112539	-----AGAGCCCTCCGGCTPAAAACTBAAAAAACATCAATCATTTATGAAAAATATGGGAA	112593
QY	901	GGTGATTTCTCTAAATTTGCTTGAAGGTGCAACATTAACAGTTTACAGGGGATTAACGTGAAT	960
Db	112594	GGTGACTACTCTTAATCTTTCTAGAGGAGACACCTTAAAGCTTGTCTCAAAATTAAGGAAGT	112653
QY	961	AGTTTTCAGGAGAGAGTGTATAGCTATATGATTTAGGAAAGAAATGGAATATACATACAGAT	1020
Db	112654	GGTTTTCAGGAGAAATTTTGGTATGATGAACAGTCTGGAGGAAAGTAGTACCTACTAAT	112713
QY	1021	GGAACTTATACTTTAATGCAATTTGAATTCCTCCAGCTGGTTATAGTATGCGACAGCCAAATC	1080
Db	112714	GGGACTATCGTTTAAAGTACCTAAACCTCCTCAAGGATATGAGAGTTGCCACACCACAT	112773
QY	1081	ACTTTTAAGTTGAAGTGGCAAACTGT--ATACTATTTATGATGGAAGAAACAGATTGAA	1137
Db	112774	ACTTTTAAAGTTGCGCTGAGGAAAGTTTTAATPAAAAAATTAAGAAAGGCTCAATTTGTAGAG	112833
QY	1138	AATCCCAATTAAGAAGATAGTAGAGCGCTTACTAGTAGAAGCAATTAATGATTTTGAAGAA	1197
Db	112834	AACCAAAATTAAGAAATGCGAGAGCATATTTCTGTACACAGCTTTCAATGACTTTGAAAG	112893
QY	1198	TTTTAGCGTTTTAAC--TACACAAAACATGCAAAAATTTTATTTATGCAAAAAAATAAAT	1254

```

Db      112894 ATGTGTAATTTATCGATTTTAAATACATATGTGTAAGTTTACTATGCAAAAAATCTAAT 112953
QY      1255 GGAAGTTACAGAGTTGCTCTATTTGCTTTAATGACAGATCTAAATCTCCACGACTCTGAA 1314
Db      112954 GGAAGTAATCAAGTGTCTACGTTCAATGCTGATTTTACACTCCACGACTGACATAT 113013
QY      1315 GATGGTGGGAACAACATGACTCCAGA---CTTTACACAGGAGAGTAATTAATACACTAT 1371
Db      113014 GATCAGGAGCAAAATATTGATTCCTGATGTCAGTGAATTAAGATTAAGATATACAT 113073
QY      1372 ATTCAGAGTCGCTGACCTCTTTAAATATATCTGTAACCAAGAGATACCGATCTGACACT 1431
Db      113074 GTTCTGCGCTATGATTTGTATTAATAATGACAGCCACCAAGAGAAAGATGCTATTTTC 113133
QY      1432 TTCTTAAACATATCAAAAAAGTAATAGAAAGGTTACAGGGGAAAAAGACACGCTATT 1491
Db      113134 TTCTTAAACATATCAAAAAATCTTGTATTAAGGCTTTAAGAAAGAGGGGGA-----T 113187
QY      1492 GAGTATAGTGGCTPACTGAGACACATTTGGCGGTACTGACTGATTAATATATAT 1551
Db      113188 ACCTATTAAGACATTTACTGAAAGCTCAGTTTAGAGCAGCAACCAATTAAGCTATTTATAC 113247
QY      1552 TTCACGTGATGCTGTAATTAAGATTAAGATAACTA-----AAAGACTATGAT 1599
Db      113248 TATACAGATAGTGTGATTTAAACAACACTTAAGACATTAATGATTAACAAGGTTATCAT 113307
QY      1600 GGTTTTGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1659
Db      113308 GGTTTTGATTAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 113367
QY      1660 CAAGTATGATTAATCCCTGACAGCTAAGCTGATGATGATGATGATGATGATGATGATGAT 1719
Db      113368 GAGGATGTCACCTTAACCATGCTCAAAATCTAGATTTCTTGTCTTAATGACGATGA 113427
QY      1720 TATCATCTCTTATTTGGAAGTCAAGTGCATCCAGAAAGATTAGTATTTATGCTATG 1779
Db      113428 TACACGACGACTTATTTGGAACGACAGATATCATCAATGATGATGATGATGATGATGATGAT 113487
QY      1780 GAGATTAATAAAG---AAGTATATCTGTAATCTAATTTAATTAATTAAGAAAAAGGCTG 1836
Db      113488 GAGGATTAACAGGCTCTTATTTCTTATTTACTCTCAAGTATTAATTTCTTAACACTGTT 113547
QY      1837 ACTGTTTACGCTGATGACAGAACTAAGATTTCCATTTTGAATGATTAATAAATAAT 1896
Db      113548 ACTGGAATTTGACAGATTAAGAAAAAGATTTAATTTAATTAATTAATTAATTTCTCT 113607
QY      1897 AAGCAAGATTTGCTTTCTCAACTGTTTAAACAGATTAACAAACCTCGAATTTAAGAT 1956
Db      113608 GATGACAAAGCTATTAAGTGAACATATCCGACAAACTCTGAGAACTCAGACTTACAGAT 113667
QY      1957 GGTAAAGCAACCATTAATTTAAACATGAGGGAAGGTTTAAACACTTCAAGGTTTACAGAA 2016
Db      113668 GGAAGAAAGCTTACCTTACATTTAAGAGATGAGAAATCATGATTTGAGGGGCTACCTTCA 113727
QY      2017 GGTATTTCTTACCTTGTCAAGAAACATTTCTGAAGCTTAAAGCTTAAAGTTAATAGC 2076
Db      113728 GGTACTCTTATGAATTTACGAAACGGGCTCTTCAATTAATGAGTAAGTGAATTAATGA 113787
QY      2077 CAAGATAGCAAAATGCTACAGATTTCAAAAACAGAAATTAACAGATTAACAGACTTACT 2136
Db      113788 AAAAATGACAGATGAAAAAGCAGAAAGCCTCAGTTAAGAAAGATGAGACTTACT 113847
QY      2137 TTTGAAATTAATAAGAGCCTGTTGTTCTTCAAGAGAGTGAATCAAAAGATCAATGCTAT 2196
Db      113848 TTTGAAACCGAAAAAGATCTTGCCCAACATCGS---TTTGACACAGATGGGCTAT 113903
QY      2197 CTAGCTTTGATGATTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2234
Db      113904 CTATCTTTGTTGTTATTAATTAATTTGCTGCTGATTTGGGTTAT 113941

```

RESULT 7

```

AEO09963
LOCUS      AEO09963      11420 bp      DNA      linear      BCT_03-APR-2002
DEFINITION Streptococcus pyogenes strain MGAS8232, section 11 of 173 of the
complete genome.
ACCESSION  AEO09963
VERSION    AEO09963.1
KEYWORDS   GI:19747394
SOURCE     Streptococcus pyogenes MGAS8232
ORGANISM   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus pyogenes MGAS8232
REFERENCE  1 (bases 1 to 11420)
AUTHORS   Smoot,J.C., Barbican,K.D., Van Gompel,J.J., Smoot,L.M.,
Chaussee,M.S., Sylva,G.L., Sturdevant,D.E., Ricklefs,S.M.,
Porcella,S.F., Parkins,L.D., Beres,S.B., Campbell,D.S., Smith,T.M.,
Zhang,Q., Kapur,V., Daly,J.A., Veasy,L.G. and Musser,J.M.
Genome sequence and comparative microarray analysis of serotype M18
group A Streptococcus strains associated with acute rheumatic fever
outbreaks
Proc. Natl. Acad. Sci. U.S.A. 99 (7), 4668-4673 (2002)
MEDLINE   21977593
PUBMED    11917108
REFERENCE  2 (bases 1 to 11420)
AUTHORS   Smoot,J.C., Barbican,K.D., Van Gompel,J.J., Smoot,L.M.,
Chaussee,M.S., Sylva,G.L., Sturdevant,D.E., Ricklefs,S.M.,
Porcella,S.F., Parkins,L.D., Beres,S.B., Campbell,D.S., Smith,T.M.,
Zhang,Q., Kapur,V., Daly,J.A., Veasy,L.G. and Musser,J.M.
Direct Submission
Submitted (31-JAN-2002) Laboratory of Human Bacterial
Pathogenesis/Rocky Mountain Laboratories/NIAID/NIH, 903 S. 4th St.,
Hamilton, MT 59840, USA
FEATURES
source
location/Qualifiers
1..11420
/organism="Streptococcus pyogenes MGAS8232"
/mol_type="genomic DNA"
/strain="MGAS8232"
/db_xref="taxon:186103"
107..502
/gene="ssb"
/note="synonym: spym18_0120"
107..502
/gene="ssb"
/note="best blastp match gbl|AAK04490.1|AEO06276.5
(AEO06276) single-strand binding protein [Lactococcus
lactis subsp. lactis]"
/codon_start=1
/transl_table=1
/product="putative single-strand binding protein"
/protein_id="AAL96931.1"
/db_xref="GI:19747395"
/transl_table=1
/translation="MYNKVIAIGRLVAPPELVKATDQKHVARLSLVNRRKNSGER
EAGFISVVGKRLAETLVSAVSKSLMSIDELRTRKDKGVHYVTEVLCOSPOL
ESRAQRARENNVNDIVDLVLEEDLTPF"
complement(757..1398)
/gene="spym18_0121"
complement(757..1398)
/gene="spym18_0121"
/note="best blastp match gbl|AAC97156.1| (U49397) unknown
[Streptococcus pyogenes]"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAL96932.1"
/db_xref="GI:19747396"
/transl_table=1
/translation="MLIVLACTIGAGKSLAALGELGTDFVEAVDNNPVLDTYYO
DPKRYAFLLQYFLNKRPSIKRAYOANNNLIDRSIFEDLFLKYNKNGVTKTELD
IXQELANMLEELBGRKRPDLITIDVSDKMLERERKGRSEVDGPNPSLEQYT
HGVHEPTWEDIVKSPKRIIDGNSDFVQNPDLATVLMIDIKELHLL"
complement(1418..2395)
/gene="spym18_0122"
complement(1418..2395)
/gene="spym18_0122"
/note="best blastp match gbl|AAC97155.1| (U49397) unknown

```


		[Streptococcus pyogenes]"
	/codon_start=1	
	/transl_table=11	
	/product="hypothetical protein"	
CDS	/protein_id="AAL96933.1"	
	/db_xref="GI:19747397"	
	/translation="MTKLNSFRIGDVEIPHRVLAIPMAGVNSAFRTIAKEFGAGLV	
	VNEMISEKGLLYNNEKTLHMLHIDENHPMSIOLFGDAGELKRAADFIOTNRKADIV	
	DINWGPCNVKVRKNAAGAKWLRDDPKYHYIKAEVTSYLDIPLYKMTGNADSLAYE	
	NALAAESAGVSALAMGRTRRQMTGTCDHETTLARVSKALITKPIFGNGVRSYDAK	
	FMLEIGVDVMTIGRAMNNPYLETQINHEFTGQELPLPFAKRLDIADHKLRLIN	
	LKGETIYRERFGLAPHYLRGTAGAAKRVASRAETIAEVALEFETVR"	
gene	complement(2382..3254)	
	/gene="spym18_0123"	
CDS	complement(2382..3254)	
	/gene="spym18_0123"	
	/note="best blastp match sp Q92B45 HSLQ_STRPY_33_KDA	
	CHAPERONIN (HEAT SHOCK PROTEIN 33 HOMOLOG) (HSP33)"	
	/codon_start=1	
	/transl_table=11	
	/product="putative HSP33"	
	/protein_id="AAL96934.1"	
	/db_xref="GI:19747398"	
	/translation="MDKIKISIAQSGAFRAVYLDSTETVALAQENHTLSSSTVALGR	
	TLIAQILANOKGDSKITVKGIDSSFGHIIISVADRGHVKGYIOMTGYDIKKTATG	
	EVLVGPNGHGFYTIIDYGTNPYTSTPLITIGEDFAFYLYLSEQNPISALGV	
	LDENDKVVAGGFVMOVLPGASEELIARYEKRLQEMPAISHLASKNHYDALIAY	
	GDEPKRLSEEPISFQDCSRERREALMLTPRKADLQAMIDEDKGAETVQFCGTRKO	
	FNSEDALINDKA"	
misc_feature	complement(3360..4000)	
	/note="nra.2 (spym18_0124): sequence confirmed upstream	
	mutation-gene fragment: best blastp match gb AAC97147.1	
	(U49397) Nra [Streptococcus pyogenes]"	
	complement(4262..4915)	
gene	/gene="nra.1"	
	/note="synonym: spym18_0125"	
CDS	complement(4262..4915)	
	/gene="nra.1"	
	/note="best blastp match gb AAC97147.1 (U49397) Nra	
	[Streptococcus pyogenes]"	
	/codon_start=1	
	/transl_table=11	
	/product="N-terminal fragment of Nra"	
	/protein_id="AAL96935.1"	
	/db_xref="GI:19747399"	
	/translation="MPYVKKKDSFLVETYLEOSIRDKSELYLLFEKSPITTFHVAK	
	OTGLTALQKYCKELDPERGNMIDITIKKGIITCCVKKPEYLLQILDYSTIILKL	
	LVFFIKNGTTSQPIKRSKRYFLUSSSAYRLRESLILKLEFGLRVSKNTIVGEYRI	
	RYLIAMLYSKFGLVLYPLDHLNQIILYRFLSQSATNLRTSPWLEEPSFYMLIALS"	
gene	5345..6919	
	/gene="cdp"	
	/note="synonym: spym18_0126"	
CDS	5345..6919	
	/gene="cdp"	
	/note="best blastp match gb AAC97148.1 (U49397) Cpa	
	[Streptococcus pyogenes]"	
	/codon_start=1	
	/transl_table=11	
	/product="putative collagen binding protein"	
	/protein_id="AAL96936.1"	
	/db_xref="GI:19747400"	
	/translation="MOKRDKTNGSGANNRRQTTIGLKYFLFEVALIGIVGESIRAF	
	GAEOSNETKKTSYILIRKVAEGDYSKLEATLKLAEISGPOESESSTSGOKLO	
	LSDGTYLLETKSPQGEIAEPIETFKYAKGLVKFGDQGFVENONKEVAPEYSVTV	
	NPDDEDSFIPKPTFTPYGKTYAKANNAGTSQVYICFNVADLHSPDSIDKSETIDPDN	
	EGKEIKTHILGADLESYANNPASTINDLSOVKKVLEGYRDSSTYANLVISVER	
	AATQALAYFTFDVNDLADYHGFGLATTEALNATKEIYAAVEDRANLNPISLDEF	
	VPSNKKYSLIGTOYHESLVLDIIRMEDKQAPITPIHKLITISKYTGTLADKKEFN	
	FEIHKSGDGAISGTYPTNSGELTVDGATFTLKQSGSLIEGSLSGVSYETKEEG	
	ASDYEVSVNKNAPDGCATKASVKEDDTIIFENRKDLVPTPTGLTTDCAITLMLLLVL	
	LGILMWLIGRKGLKND"	
gene	6912..7433	
	/gene="lepa"	

		/note="synonym: spym18_0127"
	6912..7433	
	/gene="lepa"	
	/note="best blastp match gb AAC97149.1 (U49397) Lepa	
	[Streptococcus pyogenes]"	
	/codon_start=1	
	/transl_table=11	
	/product="lepa"	
	/protein_id="AAL96937.1"	
	/db_xref="GI:19747401"	
	/translation="MTNVLRLNENPLEKAFIRLVYKISITIGFLGYILFQYIFGVMI	
	NTVMSPALSAGRGIIYRLTDRYHINDVYVEVDNLTAKGRIVAAQAGDEVSTQEG	
	LINGHPREKEVPLTYPHSSGNFPYKVPYGYFIIINDYREBRLSRYGALPINOI	
	KGKISTLRVRGI"	
gene	7454..8488	
	/gene="spym18_0128"	
CDS	7454..8488	
	/gene="spym18_0128"	
	/note="best blastp match gb AAC97152.1 (U49397) unknown	
	[Streptococcus pyogenes]"	
	/codon_start=1	
	/transl_table=11	
	/product="hypothetical protein"	
	/protein_id="AAL96938.1"	
	/db_xref="GI:19747402"	
	/translation="MKRKKLLATATLGTASLNONYKAEFAGYIDGSTLVKRT	
	FPSTYDQKVMKPADYFPRKEDADNMKGRKDLIDIPGYIDGLENTKTHVGNSTKT	
	TAKESYNPFAVYKPRGVYRYTYSVGNKAGIAYDSQMTFVYVYNNRDSGE	
	AKTIVSTEGOSDKRPVLEKNEFDITSLAKYKVTGTHGSHQSFSTLLITRECE	
	KGOVNVILQSGETKRVVIGEEYFTLKDRESVLSOLPVGIEKVEEDVTKGYTS	
	ATLQGDVTDGYNMGDSKTTDKSTDEIVVNRKDTQVPGVGVTAPFAVLSIVAIG	
	VIYITKRKA"	
gene	8489..9229	
	/gene="spym18_0129"	
CDS	8489..9229	
	/gene="spym18_0129"	
	/note="best blastp match gb AAC97152.1 (U49397) unknown	
	[Streptococcus pyogenes]"	
	/codon_start=1	
	/transl_table=11	
	/product="hypothetical protein"	
	/protein_id="AAL96939.1"	
	/db_xref="GI:19747403"	
	/translation="MISORMMMTIVGVINKAIDTLILFCLVYFLAGELMPSYHLX	
	QOADSNEKKTAAOQOPREFDLALMEDYIGMLNPGTMDYPLVQKTNLEINKA	
	VDSVAMSGSLPFDTRHNDFTDDYSILIVHMAAGNAGEIIRPKLDFEENHNKAI	
	IEIERKKLVTVTFACLTAPFQOLVFNPAITNQDQRLVYISKRSQFNPVAK	
	HHTFEVAFSTCEMFSIDNRYIVGTIOE"	
gene	9246..9833	
	/gene="spym18_0130"	
CDS	9246..9833	
	/gene="spym18_0130"	
	/note="best blastp match gb AAC97151.1 (U49397) unknown	
	[Streptococcus pyogenes]"	
	/codon_start=1	

Query Match	19.3%: Score 438.4; DB 1; Length 11420;	
Best Local Similarity	60.3%: Pred. No. 6.5e-65;	
Matches	863: Conservatave 0; Mismatches 536; Indels 33; Gaps 7;	
QY	854 CAATCAACCTCAACAGCTTCAGTACTATTAGAAAGTATGCTATAGTGATTA	913
DB	5493 CTACTGAACCTAAAAAAGCTGATCTATTATGAAAAATATGCGAAGGAC	5552
QY	914 AATTGCTTGAAGTGCAACATTCAGTTGACAGGGGATTAACGTGAATAGTTT	973
DB	5553 AACTTCTAGAGGGGCAACCTTAAAGCTTCTCAAAATTTGAAGAAGGTGTT	5612
QY	974 GAGTGTTTACAGTAATGATATTGAGAAAGATTTGACATGAGTAAGTTA	1033
DB	5613 AGAGTTTGAAGAGTACTTCAAGGACAAAATTCGAGTTGCGATGAGTACAT	5672
QY	1034 TAACTGAATTGAATTTCCAGCTGTGTATAGTATCGCAGAGCAATCACTT	1093

```

Db. 5673 TAACTGAACTAAATCACTCAAGATATGATGGCGAACCTATTACTTTAAAGTTA 5732
Qy 1094 AAGTCGCAAGCTATATCTA---TTATGATGGAAGAAACAGATTGAAATCCCAATAAG 1150
Db 5733 CAGCTGGGAAAGTTTATATAAAGGATGAAGGCTGAGTTGTAGAGAACCAATAAATAAG 5792
Qy 1151 AGATAGTAGAGCCCTTACTAGTACATATGATATGATTTGAGAAATTTGCGTTTAA 1210
Db 5793 AGGTCGAGAACCTACTCTGTATACAGCCTTAATGATTTGATGATGTGTGTTTATAA 5852
Qy 1211 CTACACAAAC-----TATGCAAAATTTTATATGCAAAATTAATAATGAGACTT 1261
Db 5853 ATCTAAAAACCTTCACTCATATGGAAGTTTACTATGCAAAATTAATAATGGAACCTA 5912
Qy 1262 CACAGGTTGCTATGCTTTATATGCAATCTAAATATCTCCACGAGCTCTGAATGCTG 1321
Db 5913 GTCAGGCTGTTTACTGTTTCAACGTAACCTTACATTCCTCCGACAGAGCTTTGACAAAG 5972
Qy 1322 GGAAGAACATGATCCAGACTTTTACACAGS---AGAAGTAATATACATCATATTTGCGAG 1378
Db 5973 GAGAAACGATGATCCAGACTTTTATGAGGGGAAGGAATTAATTAATCTACATATTTGG 6032
Qy 1379 GTCGTGACCTCTTAAATATCTGTGAACCCAGAGATACCGATCTGACACTTCTTAA 1438
Db 6033 GTGCTGACTGTTTATAGTATGCTAATAAACCTAGGGCTCAACAAGATGATATATAT 6092
Qy 1439 AACATATCAAAAAAGTAATGAGAGGTTACAGGAAAAAGCAACGATTTGAGTATA 1498
Db 6093 CTCAGATGAAAAAGTTTATGAAAAAGGTTATGCTGTGATGATCCAGCAGCTATATCGAA-- 6150
Qy 1499 GTGCTTACTGAGACCAATTTGGTGGCTACTGCAATTTGCAATATATATTTCACTG 1558
Db 6151 -----TTTAACTTCGCTGATATTTAGAGAGCTACCCATTTGCTATTTATTTATTTACAG 6206
Qy 1559 ATAGTCGTAATTTAGATTAAGATTAACATAAAAGATATCATGTTTGGAGACATGATG 1618
Db 6207 ATAGTG-----TAGACCTGATATTTAGACAGATTAATCATGTTTGGAGCTTTGACAA 6260
Qy 1619 ATAGTACTTTAGCAGTTGCTAAATCTTGTAGAAATCGCTCAAGATAG---TAATCTCTC 1675
Db 6261 CAGAGGCTTTAAATGCAACAAAGAAATAGTTGCTTATGCTGAAGACAGGCTTAATCTTC 6320
Qy 1676 CACAGCTACAGGACCTGATTTCTTATTCGATTAACATAAATATCAATCTCTATNTG 1735
Db 6321 CTAACTATTTCTAATTTGATTTATGTTCTCTAATAGTAATATATCAGCTTTAATG 6380
Qy 1736 GAACCTAGTGCATCCAGAGATTTAGTTGATTTATTTCTATGGAAGTAAAAAG--- 1792
Db 6381 GAACACAAATATCATCCAGAGCTTTAGTGATATTTATTCGAATGGAAGATTAACAGGCTC 6440
Qy 1793 AAGTATACCTGTAAGCTATATTTAATTAATGAGAAAAAGCGTACGTTTACGTGCTG 1852
Db 6441 CTATTATTCCTATTAATCAACAGTATCTAATTTCTTAAACGTTTACGGAATCTTTGAG 6500
Qy 1853 ACAGACTTAAAGATTTCCATTTTGAATTTGAATTAATAAATAATATAGCAAGATTTGCTT 1912
Db 6501 ATAGAAAAAAGATTTTAACTTTGAAATACATTTAAATCTTTCTGATGAGACACCTTAA 6560
Qy 1913 CTCAAAGCTTTAAACAGATTAACAAACCTCGAATTTTAAAGATGTAAGCAACCATTA 1972
Db 6561 GTGAGACATATCCAGAAACCTGAGACATCAGAGTATAGAGTAAGGAAAAAGCTACCTTCA 6620
Qy 1973 ATTTAAACATGGGAAAGTTTAACTTCAAGCTTTACGAAAGGTTATTTCTTACCTTG 2032
Db 6621 CATTAAAGGATGGAGATCATATGATGTGAGGGGCTACCTTCAGTTACTTATGATAA 6680
Qy 2033 TCAAGAAACAGATTTCTGAAGGCTATTAAGGTTAAATTAAGCAAGAACTAGCAATG 2092
Db 6681 TTACCGAAACGGGTGCTTCAAGATTAAGGTAAGTGTATTAAGAAAAATCACCAGATG 6740
Qy 2093 CTACAGTTTCAAAAAAGGATAACAAGTATGAGACACTTGTCTTTGAAAAATATAAAG 2152

```

```

Db 6741 GAAAGCAGCAAGAACCTCACTTAAGGAGAGCAAGAACTATTTGAAAAACCGTAAG 6800
Qy 2153 AGCGTGTGTTCACACAGACTGATCAAAAGATCAATNGCTATCTAGCTTGTATGATA 2212
Db 6801 ATCTGTCCACCAACAGCTGTTTGACAAACAGATGGGGCTATTAATCTTTGTTATATAC 6860
Qy 2213 TCGCTGATATGATTTGGGATCTGGGAAATTCACAGATTAAGATTAAGAA 2264
Db 6861 TTGTTCTACTGCGGTTATAGGTTTGGCTAATTTGGTCGTAAGGAGCTAAAAA 6912

RESULT 8
STREPTNF 2845 bp DNA linear BCT 24-NOV-1993
LOCUS streptococcus pyogenes fibronectin-binding protein (prtF) gene,
complete cds.
ACCESSION L10919.1 GI:425479
VERSION L10919
KEYWORDS adhesin; fibronectin-binding protein; protein F; prtF gene.
SOURCE Streptococcus pyogenes
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2845)
AUTHORS Sela,S., Aviv,A., Tovi,A., Burshtein,I., Caparon,M.G. and Hanski,E.
TITLE protein F: an adhesin of Streptococcus pyogenes binds fibronectin
JOURNAL Mol. Microbiol. 10 (5), 1049-1055 (1993)
MEDLINE 95020565
PUBMED 7934855
COMMENT Original
FEATURES
source: Streptococcus pyogenes (strain JRS75) DNA.
location/qualifiers
1..2845
/organism="Streptococcus pyogenes"
/mol_type="genomic DNA"
/strain="JRS75"
/db_xref="taxon:1314"
425..2404
/gene="prtF"
425..2404
/gene="prtF"
/function="binds fibronectin"
/codon_start=1
/transl_table=11
/product="protein F"
/protein_id="AA26964.1"
/db_xref="GI:425480"
/translation="MFVKGSEKMNKIFLNKEASFLAHTKRRKRPATLVGVFMYLLAC
AGAGFGOVAFAADEKTVPSHSSNPPEPWGYDAYKEYGVYIMIRYHDLRYNLG
SRQYQVCFNIOSSYPSOKNSFKTKMKKIDGNKSPDYAHTTKLKEDEKRLSL
LYNAIPDANSYKRGLEHLNATITQYAVMHTSDNSOYFETLWSEKKEGKISRQY
TLMEARKLIDPNIENATVAVNKIPSGYRLNTESENDAVYOLLSAEVYPDPPPGST
SEHNKRPEDLGPPIPDPPKHFDNLBPTLPVVLDEEVEVESLEPALPLME
LDGQEVPEKSIDLPIEVPREYFNKKQSPLAGSGSTEVITEVYGQONFVDIDKL
PNETGSGNVEETEDTKPEVLMGSGSEVEFTKDQYMGSGOTTPOVEDTKEPEV
LMGQSGSVEETKDTQMGSGOTTPOIETEDTKEPEVLMGQSGSVEFTKDQMGSG
OTTPOVEDTKEPEVLMGQSGSEVEFTKDQYMGSGOTTPOIETEDTKEPEVLMGQ
SESEVEFTKDQYMGSGSESEVEFTKEDTKEPEVLMGQSGSEVEFTKEDTKEPEV
ATGDIENYLAFLGILILSVLSIFSLNKNKVV"
repeat_unit
1327..1457
/gene="prtF"
/feature="repeat domain 1"
repeat_unit
1631..2170
/gene="prtF"
/feature="repeat domain 2"
BASE COUNT 977 a 472 c 616 g 780 t
ORIGIN
Query Match 9.58; Score 215.8; DB 1; Length 2845;
Best Local Similarity 57.28; Pred. No. 5.1e-27;
Matches 466; Conservative 0; Mismatches 322; Indels 27; Gaps 3;
Qy 51 GGTATTAAGTAAACACTCAACGATTTACTGTCACTTATAGTGGAGCTTTTAATGAT 110

```

```

Db      487 GGCACACACAAAAAGGCGATTGCTGCTACCTTAGTGGAGCTTTTATGCT 546
QY      111 CTTGCGTTGGTACTTCATGGTGTGTAGACTGTTTGGTTTACTAATCCTC 170
Db      547 TTGGGATGGCGGGGTGCTATCGGTTTGGTCAAGTAGCCTATGCTCCCATGAAGAC 606
QY      171 GAGCCCAACGCAATTAATCCAGATTCAGATTGCAATAGATGATGATGATGATG 230
Db      607 TGTGCGGAGTCATAGTAGTCCAAATCCT-----GAGTTCCCTTGGAGCGTTATGATGC 660
QY      231 TTATGTAGAGGGGCATCATATTATAACAGTTTAAAGTAGACAGCATTTAAAGGTTAA 290
Db      661 TTATGCTAAGGAATATCTCGATATATATATTGACTCGTTATCATGACTTGAAGATAA 720
QY      291 CTAGAGAGAGTAGAGATTATCACTTTATTCCTTTAATTAAAGAAACATTTCTCT 350
Db      721 TTTAAATGGAGTGGAGCTACCAAGTTTATTTTAAATTAATCAATTAATATCTAG 780
QY      351 CGGATCAGATAGTAGTGTAAAGAGTGTAAAGAAACATGATGAAATCTCTCAAAATT 410
Db      781 CCAAAAAAATAGTTTCATATAAAATTTGGTTAAAAAATTGAGGAAATGGCAATCCTT 840
QY      411 TGAAGATTATGCGATGACCCCTAGAAATTACGGAGATGAGCTAAATCAGAAATTACGAGC 470
Db      841 TGTAGATTATGCTCAGACTACTTAAGTAGTAAGAAAGAGTGGAGCAGAGCTCTTAAG 900
QY      471 TGTATGTATTAATGAGACATCCCAAAATGGCATGATGATTAATGGAAGGCTTGAACCTT 530
Db      901 TTTACTGTATATGCTATGCCATCCCAAGATGCTAATGTTATATTAAGGACTGGAGATTT 960
QY      531 GAATGCTATCAGAGTTACACAAAGAGCGGATGATGCTATTCATTAATGCTCTTATTTTC 590
Db      961 AAATGCTATTAATGCTATTAACAAATATGAGTTGGCATTTACTAGATTAATTCACAAATATCA 1020
QY      591 TAATCCAGATGAAGTTTAAAGGAGTACAGAAAGTACTGTTAGTACTTCTCAATT 650
Db      1021 ATTTGAGACATTTTGGGAGTCTGAGGCGAAAGAGTAA--GATTAGTAGTACACAAAGT 1077
QY      651 ATCTTTGAGCGCTCAAGCTTTGAAGCAATGATTCGCAATTTGGCACTAAATGCC 710
Db      1078 TACTTTAATGCGGAGGACATGAAAAACTAATTGATCCCAATTTAGAACTCTGCACT 1137
QY      711 AAAACAAGTCCCGATGATTTTACAGTAATGATTTTGAAGTCTGAGAGCAAGGAGATAA 770
Db      1138 TATTAATATCCATCAGGGGATGCTTAATATCTTTGAGTCTG----- 1181
QY      771 ATATTAATTAAGATACCAAAATCTTTGAGTGTGTTAGTTCTACTAATAACCAAC 830
Db      1182 --AAATGAAGCTTACCAAAATCTTTTAAAGTCTGAATATGATGATGATCCCTTAA 1239
QY      831 TCCAGAGACCCCAACCAATGCTCCAAATCAACTC 865
Db      1240 ACCTGGTGAAGCTCAGAACATTAATCTTAAACTC 1274

```

```

RESULT 9
SPSFB   S.pyogenes sfb gene for fibronectin-binding protein (partial).
LOCUS   X67947.1 GI:511149
VERSION X67947.1 GI:511149
KEYWORDS fibronectin-binding protein; sfb gene.
SOURCE  Streptococcus pyogenes
ORGANISM Streptococcus pyogenes
          Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
          Streptococcus.
REFERENCE 1 (bases 1 to 2263)
AUTHORS  Talay,S.R.
TITLE     Direct Submission
JOURNAL   Submitted (31-JUL-1992) S.R. Talay, GBF, Mascheroder Weg 1, 3300
          Braunschweig, FRG
REMARK   revised by [4]

```

```

REFERENCE 2 (bases 1 to 1030)
AUTHORS  Talay,S.R., Valentin-Weigand,P., Jerlstrom,P.G., Timmis,K.N. and
          Chhatwal,G.S.
TITLE     Fibronectin-binding protein of Streptococcus pyogenes: sequence of
          the binding domain involved in adherence of streptococci to
          epithelial cells
JOURNAL   Infect. Immun. 60 (9), 3837-3844 (1992)
MEDLINE   92363585
PUBMED    1386839
REFERENCE 3 (bases 1 to 2263)
AUTHORS  Talay,S.R., Valentin-Weigand,P., Timmis,K.N. and Chhatwal,G.S.
TITLE     Domain structure and conserved epitopes of sfb protein, the
          fibronectin-binding adhesin of Streptococcus pyogenes
JOURNAL   Mol. Microbiol. 13 (3), 531-539 (1994)
MEDLINE   95089690
PUBMED    7527894
REFERENCE 4 (bases 1 to 2263)
AUTHORS  Talay,S.R.
TITLE     Direct Submission
JOURNAL   Submitted (13-JUL-1994) S.R. Talay, GBF, Mascheroder Weg 1, 3300
          Braunschweig, FRG
COMMENT   On Jul 17, 1994 this sequence version replaced gi:47433.
FEATURES
     source
         1..2263
            /organism="Streptococcus pyogenes"
            /mol_type="genomic DNA"
            /strain="DSM 2071"
            /db_xref="taxon:1314"
            /db_xref="SFB"
            /gene="sfb"
            /cds
                124..2040
                /gene="sfb"
                /cds_start=1
                /product="fibronectin-binding protein"
                /protein_id="CAA48133.1"
                /db_xref="GI:511150"
                /db_xref="SPTREMBL:001924"
                /translation="MNNKMLEKNEAGFLAHTKRKREAVTLIVGFEMLASGATIGF
                GVAADERTVPHRVSONPEPFGYGDPTKPTRYHNLQNLMSKSTVOACFNKRP
                FERKRGSTFPMNTRKMDSEETFYATNPKNRDNSSVIVDELEKNTLRILYNRP
                NNGNGIMEGLEPILNALIVQNAWVYSDNSIFINDNFTTAKDLINIKPEGLSNRP
                ALKLLIDPKLSESLRPVSTRLNLFESODKLYQNLIAEYVPEPNPAPGEPBERGP
                KPELDGPTPIPEGPORNSLEPTLPVMDGOEYVPESESLPALPLMPELDGOE
                VPEVPESESLPALPLMPELDGOEYVPESESLPALPLMPELDGOEYVPESESLPALPLMPELDGOE
                YTEVYGNQONPVYDIDKKLPNENGSEGVNVEETDPEPVLMGQSESEFPRKDOTG
                MSQITPVYEIEDTKREYVLMGQSESEFPRKDOTGMSQITPVYEIEDTKREYVLM
                GQSESESEFPRKDOTGMSQITPVYEIEDTKREYVLMGQSESEFPRKDOTGMSQIT
                EVTIVEDTRPKLVFHFDDNNEPVEENREKPTKNITPIIPATGDLENVLAFLGILITS
                VLPFSLKKQTKO"
                124..270
                /gene="sfb"
BASE COUNT 790 a 383 c 477 g 613 t
ORIGIN
Query Match 7.9%; Score 180.2; DB 1; Length 2263;
Best Local Similarity 56.6%; Pred. No. 5.9e-21;
Matches 422; Conservative 0; Mismatches 278; Indels 45; Gaps 3;
QY      51 GGTATTAAAGTAAACAAACGATTACTGTCACTTTAGTGGAGCTTTTAAATGAT 110
Db      165 GGCACACACAAAAAGGCGATTGCTGCTACCTTAGTGGAGCTTTTATGCT 224
QY      111 CTTGCGTTGGTACTTCATGGTGTGTAGACTGTTTGGTTTACTAATCCTC 170
Db      225 TTTGGCATCCGCGGGGCTATCGTTTGGTCAAGTAGCCTATGCTCCGAGAGGAAC 284
QY      171 GAGCCCAACGCAATTAATCCAGATTCAGATTGCAATAGATGATGATGATGATGATG 230
Db      285 TGTGCGTCAATAGTAGTCCAAATCCT-----GAGTTCCCTTGGAGCGTTATGATGC 338
QY      231 TTATGTAGAGGGGCATCATATTATAACAGTTTAAAGTAGACAGCATTTAAAGGTTAA 290

```

Db 339 CTACAAAGGACCTTACACAGATAT-----CATTAATTACAGCTTAA 380
 QY 291 CTTAGAGAGAGTAGAGTATCAAGTTTATTCCTTATTTAAAGAAAGCATTTCTCT 350
 Db 381 TCTCAACGAGAGTAAACATATCAGCTTATGTTTATTTAAAGATTGGAACCTAA 440
 QY 351 CGGATCAGTACTAGTATTAANAAGTGTATTAANAACATGATGGAATCTCACAANTT 410
 Db 441 GAAACAGAGTATGTTTTCGAAATGGTATTAAGAGTGCGATGGTAGTGAAGACATTT 500
 QY 411 TGAAGATTATCGATGAGCCCTAGAA-----TTACGGAGATGA 449
 Db 501 TGTAAATATGCTGTATTAACCCAGAAAGATTAATGAAGTTCCAGATTAATGATGTA 560
 QY 450 GCTAAATCAGAGATTAGCAGCTGTTATGTAATGACATCCACAAATGCCAATGAT 509
 Db 561 ATTGAAAAAATATATTAAGAGTGTATTAATGATGATCCGATTAACGGTAATGAT 620
 QY 510 TATGAGGCTTGGAGCCCTGGAATGCTATGATGATTAACAGAGAGCGGCTATGCTACTA 569
 Db 621 CATGAGGAGTTAGAACATTAATGCTATCTTAATCTCAAAATGCTTTGGTACTA 680
 QY 570 TTCTGATATGCTCTTATTTCTAATCCAGATGAAGTTTAAAGGAGTCAGAAAGTAA 629
 Db 681 TTCTGATATGCTCTTATTTCTAATCTGATTAATCTTCTTACGACTGAAGTAAAGATCT 740
 QY 630 CTGCTAGTACTCTCTCAATTAATCTTGAATGCGTCAAGCTTTGAAGCAACATGATGATCC 689
 Db 741 TAAATTAACCTGAGCAATTAATCAATTAATGCTGCTGCTCAATTAAGAAATGATGATGCC 800
 QY 690 GAATTTGGCACTAAATAGCCAAACAGTCCGATGATTTTCAAGCTTAAGTATTTTGA 749
 Db 801 TAACTTATCTGAGATATTAATGAACCTGCTGATCTTCTTACGCTGAATTAATTTTGA 860
 QY 750 GCTGAGGACAGAGGAGATTAATAT 774
 Db 861 ATCTCAGATTAAGTATATCAAAAT 885
 RESULT 10
 LOCUS GGU31115 2066 bp DNA linear BCT 27-AUG-1996
 DEFINITION Group G streptococcus fibronectin binding protein (gfba) gene,
 complete cds.
 ACCESSION U31115.1 GI:950168
 VERSION U31115
 KEYWORDS
 SOURCE Streptococcus sp., 'group G',
 ORGANISM Streptococcus sp., 'group G',
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.
 REFERENCE 1 (bases 1 to 2066)
 AUTHORS Kline,J.B., Xu,S., Bisno,A.L. and Collins,C.M.
 TITLE Identification of a fibronectin-binding protein (gfba) in
 pathogenic group G streptococci
 JOURNAL Infect. Immun. 64 (6), 2122-2129 (1996)
 MEDLINE 96239026
 PUBMED 8675316
 REFERENCE 2 (bases 1 to 2066)
 AUTHORS Kline,J.B., Xu,S., Bisno,A.L. and Collins,C.M.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUL-1995) James B. Kline, Microbiology & Immunology,
 University of Miami, 1600 NW 10th Ave, Miami, FL 33136, USA
 FEATURES
 source
 1..2066
 /organism="Streptococcus sp. 'group G'"
 /mol_type="genomic DNA"
 /strain="1750"
 /db_xref="taxon:1320"
 /gene="gfba"
 /gene="gfba"
 252..1994
 252..1994
 CDS
 /gene="gfba"

/note="gfba"
 /codon_start=1
 /transl_table=11
 /product="group G streptococcal fibronectin binding
 protein"
 /protein_id="AAB06623.1"
 /db_xref="GI:950169"
 /protein_id="AAB06623.1"

BASE COUNT 724 a 326 c 435 g 581 t
 ORIGIN

Query Match 7.4%; Score 167.6; DB 1; Length 2066;
 Best Local Similarity 55.7%; Pred. No. 8.3e-19;
 Matches 398; Conservative 0; Mismatches 289; Indels 27; Gaps 3;

61 AAAAAGCTAAAGATTTACTGCTCACTTATGAGAGCTCTTTTATGATCTTCCTTGG 120
 321 AAAAAGAAAGCGATTTGCTGCTTACCTTATGAGAGCTCTTTTATGCTTTGGCATGC 380
 121 GTAAGTCCAGTGGTGGGCTGTAAGCTGTTTGGTTTATGTAATCCGCGAGCAAC 180
 381 GCGGTGCTATGCTTTGGTGGTCAAGTATGCTTATGCTGCGGATGACAGACAGCTGCGAT 440
 181 GCAATTAATCCAGATTTCAAGTTCGGAATACAGATGTAATGATTAATGATTAAGTA 240
 441 TTTAAAGGCCAGATCA-----GATTATCCCTGATGATGATGATGATGATGATGAT 494
 241 GGGCATCCATTAATTAACAGATTTAGAGTACAGACAGATTTAAGGCTTAATTAAGTA 300
 495 GCGTTTAAAGATAT-----CATGATTTAAATGATTAATTTGAAGAT 536
 301 AGTAGAAGTTATCAAGTTTATGCTTATTAATTAAGAAAGCATTTCCCTGATGATGAT 360
 537 AGCAGCGCTTACCAAGTTTATGCTTATTAATTAAGTAAGCAAGAACCGATTAAGTTAT 596
 361 AGTACTGTTAAAGAGTGTATTAAGAAACATGATGATCTTACAAATTTGAAGATTAAT 420
 597 GGTTCAGAAATATCCGTTCAAAAAGTGTATGATGATGATGATGATGATGATGATGAT 656
 421 GCGATGAGCCCTAGAGTATGAGGAGATGATGATGATGATGATGATGATGATGATGAT 480
 657 GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 716
 481 AATGAGATCCCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 717 AATGATATCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 776
 541 AGAGTTACAGAGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 777 TTAGTAACCTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 833
 601 GAAAGTTTAAAGGAGTACAGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 834 GAAAGTGAAGTCTGAATTAAGAGATGAGAGAAATTAAGAGTGAAGTGAAGTGAAGTGAAG 893
 661 GCTCAAGCTTTCAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 894 CGTGAATGCTTAGAGAAATTAATATCAGATGATGATGATGATGATGATGATGATGATGATGAT 953
 721 CCGATGATTTTACCTAAGTATTTTGAAGTGAAGCAAGGAGATTAATAT 774
 954 CCGATGATTTAATGATTAATCTTTTATTTCAAAAGATTAATCTGATCAAAAT 1007

[illegible]

OY		231	TGATGTAGAGGGCATCCATTCTTTAAACACTTTAGAGTAGCACACAGATTTAAGGGTTA	290
Db		686	GTAATCGGGAAGATTGTTAAGATATCAACAATCTCAAAGTAA-----	726
OY		291	CTTAGAAGAGAGTAGAAGTTATTCAAAGTTATTCGTTAAATTAAAGAAGCATTTCCCT	350
			+	
Db		727	-----ATGGAGTAGAAGAGATCAAGCCTATGTTTTAACCTGAAAAGTTTGAGCCAA	781
OY		351	CGGATCAGATAGTAGTGTTTAAAAAGTGATATAAAAAACATGATGAAATCTCTACAAATT	410
Db		782	AAAAGAAAGAAAGTAGCAGTCTCAATTGTTATATAAAAAATTAGATGGGATGCAGAACTTT	841
OY		411	TGAAGATTATPGCATMGAGGCCCTACAATTAGCGGAGATGAGCTAAATACGAAGTTACGAGC	470
			842 TAAAAAATAGCCTGAAAACCCTAGATTAGTGAGAGAGAAATTAAGACACATATATTGAA	901
Db		471	TGTTATGATATATGAGACATCCACAAAATGCCAAATGATATTATGGAAGCCTTGGAAACCTT	530
Db		902	GGTTCTCACAATGGATATCC---TAATAGTAAATGAATATATGAAGGGATAGACCCTCT	958
OY		531	GAATGCTATCAGAGATTACACAAGAGCGGATATGCTACTATTCTGATATATGCTCTATTC	590
Db		959	CAAGCGTATCTTAGTAACACAAAATGCCATATGGTATTAACTCGGATAGTCACCGATTAA	1018
OY		591	TAATCCAGATATAAGTTTTAAAGGAGATCAGAAAGTAACTTGGTTAGTCTTCTCAATT	650
Db		1019	TGATATTATTAATCTTCTTTACTAGTGAAGCAATGATCTTAATATACCACCTACGACGCT	1078
OY		651	ATCTTGAATGCGTCAAGGTTTGAAGCAACGTATGATCGAATTGGCAACTTAATAATGCC	710
Db		1079	AACGTTAATGCGTAGAGCTCGAAGAAAATCTAATTTCCCTGAT-----GAAAATTGGT	1132
OY		711	AAACAACTTCCGGATGATTTTCAGCTAAGATTTTTGAGCTGAGAGCAAGSAGATTA	770
Db		1133	TAAACAAGTCCATACGAACTTTAAACTTAGATTTTTTGAGTCT-----	1175
OY		771	ATATATTAAGAGATACCAAAATCTTTTGAGGTGGTGTATGTTCTACTAAACCAAC	830
Db		1176	-TCGATATAGTCCTACCAAAATCTTTTAAGTCGATATGTACTGATGATCTCTCTAA	1234
OY		831	TCGAGAGACCCACCAATGCTCTCCAATCAACCTC	865
Db		1235	ACCTGGTGATACGTACGACATATATCTTAACAATC	1269
RESULT 12				
AF009914				
LOCUS		AF009914	321 bp	DNA linear BCT 13-FEB-1998
DEFINITION			Streptococcus pyogenes fibrinectin binding protein F gene, partial cds.	
ACCESSION		AF009914		
VERSION		AF009914.1	GI:2267181	
KEYWORDS				
SOURCE				
ORGANISM				
			Streptococcus pyogenes	
			Streptococcus pyogenes	
			Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
			Streptococcus.	
REFERENCE				
AUTHORS			1 (bases 1 to 321)	
TITLE			Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.	
			protein F, a fibrinectin-binding protein of Streptococcus pyogenes,	
			also binds human fibrinogen: Isolation of the protein and mapping	
			of the binding region	
JOURNAL			Microbiology 144 (Pt 1), 119-126 (1998)	
MEDLINE			98129085	
PUBMED			9467904	
REFERENCE			2 (bases 1 to 321)	
AUTHORS			Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.	
TITLE			Direct Submission	
JOURNAL			Submitted (23-JUN-1997) Medical Microbiology, Lund University,	
			Solvegatan 23, Lund 22362, Sweden	
FEATURES			Location/Qualifiers	
source			1..321	

```

/organism="Streptococcus pyogenes"
/mol_type="genomic DNA"
/strain="M4 strain 281C"
/db_xref="taxon:1314"
>1..>313

```

```

/organism="Mycoplasma penetrans"
/mol_type="genomic DNA"
/strain="Hf-2"
/specific_host="Homo sapiens"
/db_xref="taxon:28227"
/note="clinical isolate"
104 . 1006
/gene="MYPE630"
gene

```

CDS
104. .1006
/gene="MYPE4630"
/note="paralog family"
/codon_start=1
/evidence=not_experimental
/transl_table=4
/product="hypothetical protein"
/protein_id="BAC44253.1"
/db_xref="GI:26453923"
/translation="MFKRRKINDFLNKYLSPKGEFNDNDLFDYSQWQDLEERT
KEHOKVKKRLSLNVDLKLKEVEBSKDLTKTINDPCTIPTKDKKNTIYWSIFISLI
SRLLHGFYLFALIAVYVSABHLRPVYGLMKRRLLYITTVVAVINPFLAALIF
SGFKSKVGRSIIIGIATANAIYRETKRLIKYLLDLVNSDADKENDPKIKKO
VILLIYNOYFHNLLDTPFLDELFTNDVNNLKASRSKSIINKNFKICINLYNEQTRP
LLHMFSTFYVAGIAYLYFLR"
1197. .1454
/gene="MYPE4640"
1197. .1454
/gene="MYPE4640"
/codon_start=1
/evidence=not_experimental
/transl_table=4
/product="hypothetical protein"
/protein_id="BAC44254.1"
/db_xref="GI:26453924"
/translation="MFEVNNKIKSTFOKEIKKCMWEKMFNFCNFYITSVKNNLIYE
AVHLKIANKISNCTDFLEKRSIFRFKILIGDLNA"
1447. .2370
/gene="MYPE4650"
1447. .2370
/gene="MYPE4650"
/note="MYPE4630 paralog"
/codon_start=1
/evidence=not_experimental
/transl_table=4
/product="hypothetical protein"
/protein_id="BAC44255.1"
/db_xref="GI:26453925"
/translation="MKNPNDFPKTKNFKRRKTELNGEKVYTDGLGLIAGVILKO
FESKQKRLKHLFSLNVDIKELQONNLSTINEPCIPILNNKKNKLYWSLLI
ALISRFEGHILVPTLLAVISKYLPGLSGITTNHLSREDVYVSFVCLIGNLIF
VLFTSGRANKERYRRTIISTISNIFLYRRTKILVKCYWLNSNNDQMNEDKE
LIEETKHNILINGFVFLGLTFLNNKRLNGKNDKRYKRYSLYMENTVSYLYC
FYINSIALIYFESIFIVAGFYLYQTFEQ"
2480. .3340
/gene="MYPE4660"
2480. .3340
/gene="MYPE4660"
/note="MYPE4630 paralog"
/codon_start=1
/evidence=not_experimental
/transl_table=4
/product="hypothetical protein"
/protein_id="BAC44256.1"
/db_xref="GI:26453926"
/translation="MFEVNNKSTFKKEEIKNYKDFSDRSSEHLNORIFEVOKIR
KSLINVDPEKLINEIQSKNFEOTLNEPCIPIDKKNKRIYLSIFISILRPIYVHG
KIVLCTVLLFLPKILPKYSKEMIIYLLFPLAFYLNSTICYOILYFNGKMETYKI
RGLIGESTNIEFTYGGTFRLYKSKFOLMSNNKEEDNKEFVPLINAVENYLLRTF
FSNTEPKDKOINSSTOKRTIYKIRFLPYFGILNLYPYCLVYSMLFYFCIGQII
VOILINILF"
3598. .4521
/gene="MYPE4680"
3598. .4521
/gene="MYPE4680"
/note="MYPE4630 paralog"
/codon_start=1
/evidence=not_experimental
/transl_table=4
/product="hypothetical protein"
/protein_id="BAC44257.1"
/db_xref="GI:26453927"
/translation="MLRKSKEKSKVFKTIKMKNSVDENTKVNVDLTFETWDLLEK

gene
OKTNNSHKLKFSISLNVNLSIKIKETLNFENFWTLKEPCIPILNNKKNKIYWSIFI
AIIISRFEGFYLLFAIVSLIYKLYVNSSEELVRRKMVHSIIAAYGLNPLVSG
LYWGEKKDKTFERRRTIISTNINIEPVQNTSKYKCFILSTNDIKTYTKMLKFKY
NKKLFYIYVHHHNLIDTPDEFTNOSKNVSKNRIIKPLKKEHVLISIKFY
SIYSTLIHMLISVLIYAGFYLYQFLONL"
4647. .5936
/gene="MYPE4690"
4647. .5936
/gene="MYPE4690"
/standard_name="pgl"
/codon_start=1
/evidence=not_experimental
/transl_table=4
/product="glucose-6-phosphate isomerase"
/protein_id="BAC44258.1"
/db_xref="GI:26453928"
/translation="MKLFTNPLBGNKSTKTEIKFSYASKIETLNNISIEKKCKGASM
LGMDVPSKYNKNEIKELINVSEQWNNKIRNVVVGIGGSYIGVRAIGDVLDPEN
RREKIYVSSMSYVSLIEKLKEDFYLIYSKGTLEIGVAFRLFYSLLFERFG
SEGAKERYVAITDKRGTLRKIDADIODTFSIPDPVGRFSAITPVGLFAMGVGLD
VKNVLRGCAKAIEDTRKEDPIKENTAYQVAAALRHYMSKKEKNEYCVYEDALRFETE
HLKQLFASEGEKGGILPYNCLFTTDLHLSGOFLOEGNOIFEFNCISITKNPLNVI
NEFLKDEGLGFLNGRKLDYINKVAASSTIDHHIDKIDITIIQINDNNESFGILY
SMFSKVVAMSGLLKVPDPQGVGEVAKSMFTKLK"
6048. .7133
/gene="MYPE4700"
6048. .7133
/gene="MYPE4700"
/codon_start=1
/evidence=not_experimental
/transl_table=4
/product="conserved hypothetical protein"
/protein_id="BAC44259.1"
/db_xref="GI:26453929"
/translation="MYSFNNKKEKKKNPPREDFNGLNIAIVNLYIYAVLLIS
VSIYIGLISFTKPKIKAKTSFYLYAFSSAVFLMGTVGLIEGEGTETPHEIDNS
EGINOLIMAGIIGAGLALGVIAVFRYFVFESEVHTHDNHSDDHIFMLSDIN
PKAAMLVILILSHRTIDGFGVYAKASTGADNLGLFIFLIIEVEVLIYVRQ
VOYQGRKALYINFTVLAIVPMVIGVINYQVLAOVAMILPEVNSGSGSIARVQ
TELVPRLHYKMSKSDWKYLIIYCSGVIVFALFILSPTHHVEHNDHDEHTTQ
YKLIANPKQYISSLNFNIVANNPQV"
7204. .8640
/gene="MYPE4710"
7204. .8640
/gene="MYPE4710"
/standard_name="gap"
/codon_start=1
/evidence=not_experimental
/transl_table=4
/product="NADP-dependent glyceraldehyde-3-phosphate
dehydrogenase"
/protein_id="BAC44260.1"
/db_xref="GI:26453930"
/translation="MNTFYKKSATINCKFVNSEKTIPIYSTIDSTIGSVAALSKD
IDMAFEGALNLMNRLTPEMRIKIKIEFAYFLAEKEFLTMSYEIGKSYNALK
EVERSEENIEFTIEVYEKEFVPSVIGEPVNIKGTGTFYVPGVVALAPFVYPI
NGLKAIITFLVGNITVFKPAQCSLSSQAKFEDANFAGVFNVTGSGSEIGD
YLENNRIGOGAFTGSSDIEGFLAKLPKPYLVELEGDANLIVNMADVLAKEII
KCAPYSGORCAIRKIVLREKIDADDLIHLIKENOKIKVGNPLNNDPTPLIDKRAV
DNPIDLIDAIHGANLYVGCNIEBGNLNRHLIDNVNSKSLAMEEPPGRLPIIRV
STINERIALANSEYGLGVSFTKDLSEARTAKYLDGTVINNGSSRGDPIDFPFIG
IKKSGFVGQIRDSLAMTRVGIENE"
complement(8788. .9687)
/gene="MYPE4720"
complement(8788. .9687)
/gene="MYPE4720"
/note="paralog family"
/codon_start=1
/evidence=not_experimental
/transl_table=4
/product="oxido-reductase-aldol/keto reductase family"
/protein_id="BAC44261.1"
/db_xref="GI:26453931"

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 97678: contig of 97678 bp in length
* 97679 97778: gap of 100 bp
* 97779 101129: contig of 3351 bp in length
* 101130 101229: gap of 100 bp
* 101230 145265: contig of 44036 bp in length.

FEATURES

source

1. 145265
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_1lb="CH211-266K22"
/clone_1lb="CHORI-211"

misc_feature

1. 97678
/note="assembly-fragment:01313
fragment_chain:1"

misc_feature

97779..101129
/note="assembly-fragment:00879
fragment_chain:1"

misc_feature

101230..145265
/note="assembly-fragment:00921
fragment_chain:1
clone_end:T7
vector_side:right"

BASE COUNT 44802 a 27655 c 28128 g 44480 t 200 others
ORIGIN

Query Match

3.2%; Score 73; DB 2; Length 145265;
Best Local Similarity 43.3%; Pred. No. 0.003;

Matches 340; Conservative 0; Mismatches 445; Indels 0; Gaps 0;

```
QY 1323 GAAACAAGTACCTCCAGACTTTACACAGAGAGTAATAATACACTCATATTGACAGTCG 1382
DB 32864 GATTTTAAATTTCTTTAAAGTAACTAATTAATTAATTAATTAATTAATTAATTA 32923
QY 1383 TGACCTCTTAAATATACGTGTGAAACCAAGATCCGACACTTCTTAAACA 1442
DB 32924 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 32983
QY 1443 TATCAAAAAAGTAATGAGAAAGGTTACAGGAAAAAGCAAGCTATGAGTATAGTGG 1502
DB 32984 TAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 33043
QY 1503 TCTAAGTACACACATTTGGCGTCTACTCAGTTAGCAATATATTATTCTCAGTAG 1562
DB 33044 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 33103
QY 1563 TCGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1622
DB 33104 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 33163
QY 1623 TACTTAGAGTTGCTAAATCTTGTAGATACGCTCAAGATAGTATATCTCCACAGCT 1682
DB 33164 TAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 33223
QY 1683 AACTGACCTTGATTTCTTATTCGATTAACAATTAATTAATTAATTAATTAATTA 1742
DB 33224 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 33283
QY 1743 GTGCGATCCAGAGATTTAGTATTTATTCGTAGAAAGATTAATTAATTAATTAAT 1802
DB 33284 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 33343
QY 1803 TGTAACTCATTAATTTACATTGAGAAAAAGGCTGACTGTTAGCTGGTGACAGACTAA 1862
DB 33344 TATTAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 33403
QY 1863 AGATTTCATTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1922
```

```
DB 33404 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 33463
QY 1923 TAAACAGATTAACAAACCTCGAATTTAAGTGTAAAGCAACCATTAATTAATTA 1982
DB 33464 TAATGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 33523
QY 1983 TGGGAAAGTTTAACTTCAAGTTTACAGAGGTTATTTCTTACCTTGTCAAAGAAC 2042
DB 33524 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 33583
QY 2043 AGATTTCGAGGCTATTAAGTTAATTAATTAATTAATTAATTAATTAATTAATTA 2102
DB 33584 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 33643
QY 2103 AAAAA 2107
DB 33644 TAATTA 33648
```

Search completed: August 19, 2003, 12:03:45
Job time : 8158 secs